7911150

GRANT, ROSALIE MAY
GROUP AND INDIVIDUAL PROBLEM SOLVING: HIGH SCHOOL STUDENTS.

THE UNIVERSITY OF OKLAHOMA, PH.D., 1978

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

GROUP AND INDIVIDUAL PROBLEM SOLVING

HIGH SCHOOL STUDENTS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

By

ROSALIE MAY GRANT

Norman, Oklahoma

1978

GROUP AND INDIVIDUAL PROBLEM SOLVING

HIGH SCHOOL STUDENTS

APPROVED BY

John W. Renner

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

My thanks are extended to the members of my doctoral committee, Dr. Larry Toothaker, Dr. Lloyd Williams, and Dr. Guido Chincarini for their time and valuable assistance. To Dr. Tom Rhymes, thank you for your patience and help in preparing and running the computer programs.

I am particularly grateful to Dr. John Renner, Chairman of the Committee, for his direction during the preparation of this dissertation and also for his guidance throughout all of my graduate studies. By living through his philosophy of education, Dr. Renner has contributed so much to both my intellectual and personal growth. Thank you, Dr. Renner, for the freedom you have enabled me to obtain.

Appreciation is also extended to the teachers and students in the high schools which were used in this study. Without their co-operation and assistance, the data for this dissertation could not have been obtained. For their valuable assistance in preparing the scoring hierarchies, Robert Bibens, Michael Renner and Bill Fix, a sincere thank you is extended.

To my family, who first gave me the opportunity to further my education, thank you.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table | Page

LIST OF GRAPHS

GROUP AND INDIVIDUAL PROBLEM SOLVING

HIGH SCHOOL STUDENTS

CHAPTER I

INTRODUCTION

Statement of the Problem

The primary objective of this study was to compare the written responses to two science problems given by individual high school students with responses to the same problems given by groups of four students. To accomplish this objective, individual and group responses were categorized to represent different qualities of responses. Scores were allocated to those different categories. The scores received by individual students were compared with those received by groups of students.

Four factors which were hypothesized to be related to the type of response given by students to the problems were also investigated. These factors were previous experience working in groups, degree of intellectual heterogeneity within the group, level of intellectual development of students and grade level.

This study sought answers to the following questions:

1. Do the responses to two science problems by high school students who solve the problems in groups have a higher quality* than the responses by students who solve the same problems individually?
2. Do high school students who have a higher level of intellectual development provide higher quality solutions to two science problems? The level of intellectual development is determined by tests derived from Piagetian theory.
3. Do the responses of high school students in the senior grades have a higher quality than the responses to two science problems of students in the lower grades?
4. Do intellectually homogeneous groups of high school students provide solutions of a different quality to two science problems than intellectually heterogeneous groups?
5. Do groups of tenth grade students who are regularly studying a biology course which emphasizes group work provide higher quality responses to two science problems than groups of tenth grade students who are studying a biology course which does not emphasize group work?

---

*The use of the concept of quality implies that it is possible to construct a hierarchy of responses which represent qualitatively different categories of thought.

Discussion

For many years teachers have been interested in grouping students. Several possible reasons exist for this interest. Teachers hope to have more manageable classes; they want to help certain students and hope that student groups will be more productive and learn more than students working individually. Furthermore, a group may be used to help a student learn things which society requires him to learn even though the child has not yet felt any particular need to learn those things (Thelen, 1960).

In contrast to these practical, pedagogical arguments Dewey (1915) presents philosophical reasons for grouping students. He states that schools are the agencies by which society's accomplishments are presented for use by future members of society. In other words, the school should reflect the active communal spirit of society and should not be set apart from it. Schools should be organized on a social basis to form a miniature community, an embryonic society. If a student is a member of a small community, argues Dewey, which she/he freely serves and has the capability of self-direction within, the society has the most effective means of becoming worthy and harmonious. An implication of Dewey's arguments is that students should cooperate, help and communicate freely with one another. Consequently, students should be actively engaged in group work.

Psychological reasons for grouping students can be derived from the work of Jean Piaget (1973a, pp. 156-166). Social interaction and transmission with other persons, peers or adults, is one of the fundamental factors by which the individual moves from stage to stage, according to Piaget's mental-development model. Piaget states that three media of social life affect intelligence--language, the content of the interaction and the logical rules imposed on thought (Piaget, 1973a, pp. 156-166).

Therefore, within the theoretical framework of Piagetian developmental psychology, social interaction incorporating the three media suggested by Piaget should play a significant part in the classroom. Children should discuss, share experiences and argue with one another. This implies that to promote intellectual development students should work in groups. In fact, Piaget urges educators to carefully consider the activity and grouping of students when implementing the new methods of physics teaching which require student experimentation (Piaget, 1973b, p. 20).

In addition to social interaction, Piagetian cognitive theory postulates that equilibration--a self-regulatory mental process--leads to intellectual development (Piaget and Inhelder, 1963, pp. 156-159). Initially extant patterns of reasoning guide the individual's interactions with the environment. However, eventually the individual encounters contradictory situations in which the reasoning patterns are inadequate to guide behavior or explain observations made. These situations produce a state of disequilibrium or cognitive conflict which signifies that extant reasoning patterns must be changed. Provided the discrepancy between these reasoning patterns and the new patterns required to remove the contradiction is not too great, the individual may alter existing reasoning patterns. This self-regulatory process is known as <u>equilibration</u>. The new, more complex reasoning patterns not only incorporate the old patterns but also refine and extend them. This increasing differentiation and extension of reasoning patterns is characteristic of mental development. The equilibration theory emphasizes that self-regulatory processes are the basis for learning (Ginsburg and Opper, 1969, p. 228).

Piaget's equilibration theory has implications for education. If a teacher is concerned with promoting the intellectual development of students, situations should be arranged which induce cognitive conflict in students. Problem solving activities may be used to induce cognitive conflict. In fact, Lawson and Wollman (1975, pp. 470-472) have suggested that physics homework problems which meet certain criteria can be devised to arouse cognitive conflict.

One further factor in mental growth, according to Piaget, is maturation (Piaget and Inhelder, 1963, p. 154). This is organic growth, especially of the nervous and endocrine systems. Maturation opens up new possibilities for mental development and thus, is a necessary but not sufficient condition, for the appearance of new reasoning patterns. Other conditions such as cognitive conflict, social interaction and experiences gained from teaching with the environment are required before the potential for mental growth can be fulfilled.

Piagetian theory describes how the environment operates when it influences the student. The environment does not simply impose its conditions on the behavior of the student. Instead the learners are active because they interpret the events in the environment. It is this interpretation, not the events themselves, to which learners respond. Therefore, "the child modifies raw experience as much as it changes him" (Ginsburg and Opper, 1969, p. 70).

There are two types of mental experience--physical and logico-mathematical. Physical experience involves actions which extract the physical properties of objects. Knowledge such as the hardness of objects is drawn directly from the objects themselves. On the other hand, logico-mathematical experience results in knowledge which is not the direct result of perceiving the objects but of the mental operations performed on the objects (Ginsburg and Opper, 1969, p. 171). The child constructs relationships from his mental actions on the objects.

Physical experiences are essential for the very young child because it is through these experiences that mental structures about the environment are built. Logico-mathematical experiences are possible only when the child possesses the requisite mental operations so that learning can occur from using the mental operations, rather than relying only on the information obtained directly from the objects themselves. To promote intellectual development, the student must be allowed to interact with the objects within the environment.

Piaget, therefore, postulates four factors, social interaction, experience, cognitive conflict in conjunction with self-regulation, and maturation which promote intellectual development. If problem solving induces cognitive conflict, then do students, who socially interact with one another and who may therefore be in situations which provide greater opportunity for cognitive conflict, give different responses to specific problems than do students who have not had this opportunity? The present study investigated whether the quality of the responses given by groups of students differed from the quality of the responses given by individual students to the same problems.

What type of problems should students be asked to solve? When commenting on research studies which compared individual and group problem solving, Thelen stated that ". . . what is generally reported as educational achievement tends to be limited to memory of information, conditional skills in problem solving of specified types and more or less subjective opinion as to how "good" a student or citizen the child is in class" (Thelen, 1960, p. 131). On the other hand, consider the recommendations of the Educational Policies Commission. The Commission stated that the ability to think, the aim of education, comprises "--(the) processes of recalling and imagining, classifying and generalizing, comparing and evaluating, analyzing and synthesizing, and deducing and inferring" (Educational Policies Commission, 1961). These processes are known as the rational powers. Therefore, it seems unlikely that the problem solving discussed by Thelen would reflect the development of the rational powers. Hence, problems given to students in this study required the rational powers. Furthermore, Piaget recommends that students carry out their investigations (Piaget, 1973b, p. 35). The problems also required students to design and discuss their own experiments.

Definitions

Piaget's theory of intellectual development provided the theoretical framework in which to compare group and individual problem solving. The following definitions are taken from Piagetian theory.

(a) Action. An action is a response of an individual whether it be "directed toward the outside world or as an act internalized in thought" (Piaget, 1973a, p. 4). These responses are of a functional nature.

(b) Operation. An operation is a mental action that can be internalized and is reversible. It can be carried out in thought as well as in action and it can take place in both directions (Piaget, 1970, pp. 21-22).

(c) Structure. A structure is a group of logically related operations (Bautista, 1974). Mental structures serve to guide mental operations.

(d) Formal Operational Thinking. Formal operational thinking is characterized by the ability to use hypothetical reasoning based on the logic of all possible combinations and to perform controlled experiments (Inhelder and Piaget, 1958, p. xiii). One of the most fundamental properties of formal thought is the subordination of reality to possibility.

(e) Transitional Thinking. Formal thinking does not suddenly emerge from concrete operational thought. As the concrete operational student becomes better able to organize and structure information and problems, the inadequacies of concrete operational thought are recognized (Flavell, 1963, p. 209). Whilst researching for new methods to overcome these inadequacies, new mental operations are developed. For example, after dealing with a number of qualitatively different factors such as length, thickness and weight, the concrete operational student realizes that these factors are related. It is recognized that one particular effect may result from several causes. The concrete operational student gradually develops an awareness of the interaction of many variables, and new mental operations emerge from such complex situations (Inhelder and Piaget, 1958, p. 282).

Hence, the transition from concrete operational to formal thinking is not discontinuous. Many intermediate steps are involved (Inhelder and Piaget, 1958, p. 280). This gradual progression from concrete operational to formal thinking is defined in this study as transitional thinking.

Piaget has not defined a level of intellectual development known as transitional thinking. McKinnon (1971) is the earliest reference to particular qualities of thought exhibited by students moving from the concrete stage to the formal stage of thought. In McKinnon's study transitional thinking was referred to as post-concrete thinking. The definition used in the present study is a compilation of McKinnon's research findings and Piaget's mental-development model. A description of the stages of intellectual development proposed by Piaget has not been included in this dissertation. There are many discussions of these stages of intellectual development (Phillips, 1975; Ginsburg and Opper, 1969; Piaget and Inhelder, 1969).

(f) Concrete Operational Thinking. Concrete operational thinking is based upon the use of what Piaget calls concrete operations. These mental operations are termed concrete "because they operate on objects and not yet in verbally expressed hypotheses" (Piaget, 1964, p. 179).

(g) Equilibration and Disequilibrium. Whenever a contradictory situation arises in which extant reasoning patterns are inadequate, cognitive conflict or disequilibrium arises which signifies that the reasoning patterns must be changed to accommodate the contradictory situation. If the discrepancy between the new reasoning patterns required to remove the contradiction and the old reasoning patterns is not too great, the individual may alter the existing reasoning patterns. This self-regulatory process by which intellectual development occurs is called equilibration.

(h) Social Interaction. Social interaction incorporates the verbal and non-verbal communication between people in which both the recipient and the initiator participate. It is partly through this interchange of ideas that each person grows intellectually.

(i) Maturation. Maturation is the increase in age of the subject and its effect on the endocrine and nervous systems.

The following definitions were proposed for the purposes of this study.

(a) Problem Solving. Problem solving requires the student to use those mental processes which are essential to thinking. These mental processes or rational powers have been identified by the Educational Policies Commission. These processes are recalling and imagining, classifying and generalizing, comparing and evaluating, analyzing and synthesizing, and deducing and referring (Educational Policies Commission, 1961, p. 5).

The problems were devised so that they could not be satisfactorily solved by the student just recalling factual information learned by rote without understanding. Because both problems required the students to devise an experiment to solve a problem, an adequate answer would involve the use of the rational powers of analyzing and synthesizing. For this study problem solving required free-response, written answers.

(b) Group. A group comprises four students from the same school grade.

Premises of the Study

The premise for this research was that the Piagetian task interviews could be used to measure formal and concrete levels of intellectual development. Furthermore, the conclusion of the research by the Cognitive Analysis Project (CAP) (Renner, 1977) was accepted as a premise. The CAP developed incidents which measured the levels of intellectual development identified by Piaget. This conclusion was accepted as valid. It was also accepted as a premise of the study that the group responses to the problems represented the coordinated and unified effort of the group as a whole. At the beginning of the group problem-solving sessions, the group was requested to provide one solution agreed upon by all the members of the group. In two cases the response submitted by the group had obviously been written by two individuals who had widely differing solutions to the problems. Data from these groups were omitted from the study.

Limitations

The limitations of this study are related to the students selected and to the nature of the problems.

Only students in grades ten, eleven and twelve were used in the study. Generalizations cannot be made to other grades. Furthermore , only science problems were solved by the students. Specifically the content was biology and geology. It is not known how the results of this study would be affected if the problem solving involved different subject areas. A hierarchy of responses was developed for each of the problems that students were asked to solve. A limitation of the study is the degree to which the categories identified amongst the responses do, in fact, form a hierarchy which represents different qualities in the solutions to the problems.

References

Bautista, L. B. "The Relationship between Intellectual Levels and Achievement in the Comparison of Concepts Classified According to a Scheme Derived from the Piagetian Model." Ph.D. dissertation, The University of Oklahoma, 1974.

Dewey, J. The School and Society. rev. ed. Chicago: University of Chicago Press, 1915.

Educational Policies Commission. The Central Purpose of American Education. Washington, D.C.: National Education Association, 1961.

Flavell, J. H. The Developmental Psychology of Jean Piaget. Princeton, New Jersey: D. Van Nostrand Co., Inc., 1963.

Ginsburg, H., and Opper, S. Piaget's Theory of Intellectual Development An Introduction. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1969.

Inhelder, B., and Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. transl. Parsons, A., and Milgram, S. New York: Basic Books, Inc., 1958.

Lawson, A. E., and Wollman, W. T. Physics Problems and the Process of Self-Regulation. The Physics Teacher, 1975, 13, No. 8.

McKinnon, J. W., and Renner, J. W. Are Colleges Concerned with Intellectual Development? American Journal of Physics, 1971, 39, No. 9.

Phillips, J. L. The Origins of Intellect Piaget's Theory. 2nd ed. San Francisco: W. H. Freeman and Co., 1975.

Piaget, J. Development and Learning. Journal of Research in Science Teaching, 1964, 2, No. 3.

________. Genetic Epistemology. transl. Duckworth, E. New York: W. W. Norton and Co., Inc., 1970.

________. The Psychology of Intelligence. Totowa: Littlefield, Adams and Co., 1973.

________. To Understand Is to Invent--The Future of Education. New York: Grossman Publishers, Viking Press, 1973b.

Piaget, J., and Inhelder, B. The Psychology of the Child. New York: Basic Books, Inc., 1969.

Renner, J. W. Evaluating Intellectual Development Using Written Responses to Selected Science Problems. The Story of the Cognitive Analysis Project. Norman: The University of Oklahoma, 1977.

Thelen, H. A. Education and the Human Quest. New York: Harper and Brothers Publishers, 1960.

CHAPTER II

LITERATURE REVIEW

Different types of research were reviewed. Some research compared individual and group problem solving. Other research not only compared groups with individuals but also investigated how the mental ability of the individual was related to individual or group problem solving. Mental ability in those studies was measured by a general test of intelligence. A third type of study compared individuals with groups and investigated the relationship between mental ability and problem solving. In these studies, however, mental ability was determined by a specific content test in a particular subject. Criteria were developed by which studies relevant to the proposed study were selected for review. These criteria are outlined below.

Criteria

Basically the present research was concerned with the quality of the free-response written solutions to two science problems provided by individuals and groups of tenth, eleventh and twelfth grade students. Therefore, the first criterion for including a study in this review was that it must be an experimental, not philosophical, study of group versus individual problem solving in school students. Studies concerned with the affect of the group on the individual were omitted. Ideally the problems should require free-response, written solutions. This second criterion, however, was not always met because sufficient details of the cognitive measures used in a study were not reported. A third criterion dealt with the nature of the response given to the task required of the students in the research. The correct solution to the problem should not just require recall, but should also involve the higher intellectual processes, such as analyzing or synthesizing. It was not always possible to judge from the report whether this criterion was met. Brainstorming problems, on the other hand, in which groups or individuals were judged on their ability to create a large number of diverse ideas, not necessarily logically connected to the problem, were easily identified and were excluded. That type of problem was qualitatively different from those used in this study. If the studies used manipulative tasks, required oral solutions or dealt with computer-assisted instruction then those studies were omitted.

Although the Piagetian model of intellectual development is the theoretical basis for this study, the research reviewed in this section may not necessarily have that particular theoretical foundation. Some research projects pre-tested students on their problem-solving ability and then formed different types of groups according to that ability. In the present study students were grouped according to their level of intellectual development as measured by Piagetian tests. Only one study (Silverman and Stone, 1972) was found in which students were grouped according to their Piagetian level of intellectual development. This study, however, was not reviewed because the problem presented to the third grade students was the Piagetian conservation of area task which required oral responses. If a study investigated the relationship between mental ability and problem solving in school students with respect to free-response written solutions, but the mental ability of the students was determined by some measure other than Piagetian tests, then that study was still reviewed.

Group Versus Individual Studies

A study by Hug (1971) compared individual and group performances. Cognitive and affective gains of 136 biology high school students under three different pedagogical methods were measured. Students were randomly allocated to independent study, small discussion groups of four to five students, large lecture-demonstration groups of approximately fifty students or a control group which received a combination of all three teaching procedures. After seven days cognitive and affective tests were administered. No information was given in the report concerning the nature of the cognitive measure. Either free-response or multiple-choice questions could have been used. All students performed equally well on the cognitive test, irrespective of the method of presentation of the work.

Group Versus Individual-Specific Mental Ability

A variety of problems requiring manipulative, cognitive or a combination of both were given to fifteen to seventeen year old students, in a study discussed by Bos (1937). Some of the cognitive problems were free-response questions. Initially twenty-two students worked the problems individually and, on the basis of those results, students were divided into three levels of ability. Groups of two were formed with equal or unequal ability. The results showed that students, no matter how they were grouped, achieved more in groups than individually. Additional problems, one of which did not require a free-response solution, were given to the students. Again, in all cases, group performance was superior to individual performance. In fact, some groups had success when individuals were not capable of solving the problem.

The Arithmetic Reasoning section of the Stanford Achievement Test and the Sequential Tests of Educational Progress (STEP), Social Studies, Part I, Form 3A were the mental ability measures used by Hudgins and Smith (1966) to classify 152 fifth and eighth grade students. A total of forty-eight groups containing three members each were formed. Half of the groups were drawn from schools in middle-class neighborhoods and half were drawn from schools in lower-class areas. The aim of this study was to investigate the productivity of the group in relation to the most able member of the group and how peers perceived that member's ability. Each group contained one high ability and two low ability students. In half the groups, the high ability student was perceived as such by peers. This student was designated as a high-status, high-ability student. In the other half, the high ability of one student was not perceived by the other group members. This student was referred to as a low-status, high-ability student. Alternative forms of the Stanford and STEP tests were given to the groups. It was not reported whether the social studies and arithmetic questions were free-response or multiple-choice. Presumably written solutions were required.

In both the high and low socio-economic schools, the productivity on the arithmetic problems of the high-status, high-ability students who answered the problem individually was equivalent to that of the group. However, in the high socio-economic schools, the group performance on the arithmetic problems was greater than the individual performance of the low-status, high-ability member. Nevertheless, in the low socioeconomic school, the group performance on the arithmetic problem was equivalent to the individual performance of the low-status, high-ability member.

Interesting results were obtained with social studies problems. Regardless of the socioeconomic background of the students and of the perceived status of the high-ability member, the group scores were lower than the individual scores of the high-ability member. The differences between the two scores were, however, not statistically significant. The authors hypothesized that the superiority of the individuals might be due to the inherent characteristics of the social studies questions.

Further hypotheses were proposed and investigated. These hypotheses were contingent upon the outcome of the investigation with the low-status, high-ability students. It was proposed that if the group was more productive than the low-status, high-ability member, the status of that high-ability member would not change after the group solved the problems. Alternatively, if there was no difference between the group productivity and that of the low-status, high-ability member, the status of this member would change as a result of the group activity. The data supported these hypotheses. The low-status of the high-ability student did not change in the high socioeconomic schools solving the arithmetic problems. This was the group in which the group productivity was greater than the high-ability, low-status individual. The status of the low-status, high-ability student did change for these groups in which there was no difference between the productivity of the group and the low-status, high-ability individual. These groups were from the high and low socioeconomic schools solving social studies problems and the low socioeconomic schools solving the arithmetic problems.

Group Versus Individual-General Mental Ability

Two studies which used a general measure of ability for grouping students were made by Amaria, Biran and Leith (1969) and Gabel and Herron (1977). In the former study, individuals and groups of two students worked through a programmed instruction science unit. Pilot studies were carried out with ten and twelve year olds. The main study involved 277 seventh and eighth grade students. Performances on post and transfer tests by individuals and intellectually heterogeneous and homogeneous groups were the dependent variables. Transfer tests used questions which required the student to apply the concepts to new situations. Groups comprised above and below median I.Q. students. Although the I.Q. measure used in the pilot studies was not specified, the Raven Progressive Matrices test was used in the main study. Students not only had to answer the problems correctly, but also had to explain how they solved the problems. Therefore, presumably students were at least sometimes involved in providing free-response, written answers. Cooperative groups obtained higher scores than individuals, especially the lower ability students. Boys and girls performed differently when pairs were made up of intellectually homogeneous and heterogeneous groups. At both levels of intelligence, girls achieved higher scores in heterogeneous groups. Boys, on the other hand performed better in intellectually homogeneous groups except at one comprehensive school.

One aspect of the study by Gabel and Herron (1977) was to investigate the performances of individual and intellectually homogeneous and heterogeneous groups of two students with respect to learning rate, retention and attitude. An Intermediate Science Curriculum Study unit was studied by 1022 seventh grade students. The Otis-Lennon Test of Mental Ability, Form J was used to categorize students. Student achievement was determined on chapter and unit tests which contained multiple-choice, free-response and laboratory performance questions. Contrary to the Amaria, Biran and Leith (1969) study, Gabel and Herron (1977) did not find differences in the performances of intellectually homogeneous and heterogeneous groups. The learning rate of groups differed from that of individuals according to whether students attended county or city schools. County schools working individually had a higher learning rate than groups of two students. The reverse result was obtained for city students. Retention, on the other hand, for county students was less for individuals than for groups. There was no difference between individual and group retention for city students. A statistically significant interaction at $\alpha = 0.02$ occurred between mental ability and grouping. For both county and city schools, students of low ability obtained higher scores when working in groups than when working alone. Different results were obtained for high ability students. There was no difference between individual and group performances by county students, but high ability city students performed better when working alone.

Need for the Study

Several science curricula such as the Science Curriculum Improvement Study (1970) and Science 5/13 (Schools Council, 1971) have been based on Piaget's mental-development model. These curricula aim to develop the student's intellect by requiring the student to solve problems. The studies reviewed in the previous section investigated individual and group problem solving. Some studies included the factor of mental ability. That factor, however, was not measured by Piagetian task interviews. There is, therefore, a need to investigate students solving problems in groups and individually using the measures of intellectual development proposed by Piagetian theory.

Furthermore, the curricula formulated within a Piagetian theoretical framework emphasize students experimenting and solving problems in groups. Individualized instruction is not part of these curricula. There are, however, other science curricula such as the Intermediate Science Curriculum Study (1971) and the Individualized Science Instructional System (Burkman, 1974, pp. 30-32) in which the students work individually. Although these curricula are not based on Piagetian theory, they nevertheless aim to promote the ability to think through problem solving. As yet no research has been carried out to compare the responses of individuals and groups of tenth, eleventh and twelfth grade students solving science problems which require free-response written solutions and the use of the higher intellectual processes such as analyzing and synthesizing.

Frequently teachers are told that it is a "good idea" to encourage students to work in groups. Although this may be intuitively appealing to the teacher, no rationale is given for implementing group work in the classroom. The teacher then initiates group work, but has little understanding of the affects of group work on the students' learning.

Sometimes group work is proposed as a solution to the managerial problems of a teacher. Ideally, the teacher is told, each student should receive individual instruction because each student has a set of unique needs and problems. It is not feasible, however, for the teacher to provide twenty to thirty individual courses of study for each student in the class. Group work is suggested as the solution to the teacher's problem. The teacher can then provide one set of instructions suitable for all members of the group. Perhaps only six sets of instruction would have to be prepared for a class of twenty to thirty students. In both of the above examples in which group work is advocated, it is proposed as being advantageous to the teacher. Research needs to be carried out which investigates the educational benefits of group work for the students.

References

Amaria, R. P., Biran, L. A., and Leith, G. O. M. Individual Versus Co-operative Learning I: Influence of Intelligence and Sex. Educational Research, 1969, 11, No. 2.

Bos, M. C. Experimental Studies of Productive Collaboration. Acta Psychologica, 1937, 3.

Burkman, E. Unlooking High School Science: The ISIS Project. The Science Teacher, 41, No. 7.

Gabel, D., and Herron, J. D. The Effects of Grouping and Pacing on Learning Rate, Attitude, and Retention in ISCS Classrooms. Journal of Research in Science Teaching, 1977, 14, No. 5.

Hudgins, B. B., and Smith, L. M. Group Structure and Productivity in Problem-Solving. Journal of Educational Psychology, 1966, 57, No. 5.

Hug, W. E. Comparison of Cognitive and Affective Gains Between Independent Study, Small-Group Discussion and Large-Group Presentation in High School Biology. Science Education, 1971, 55, No. 2.

Intermediate Science Curriculum Study. Probing the Natural World. Morristown, New Jersey: Silver Burdett Co., 1971.

Schools Council. Science 5/13. London: Macdonald Educational, 1971.

Science Curriculum Improvement Study. Chicago: Rand McNally and Co., 1970.

Silverman, I. W., and Stone, J. Modifying Cognitive Functioning Through Participation in a Problem-Solving Group. Journal of Educational Psychology, 1972, 63, No. 3.

CHAPTER III

DESIGN AND PROCEDURES OF THE RESEARCH

The present study was primarily concerned with comparing groups and individuals solving problems. The research also investigated the student factors of previous group experience, grade, and level of intellectual development as well as degree of intellectual heterogeneity within the groups. The design and procedures adopted for this study will be described in this chapter.

Populations and Samples

Students and Schools

Data for the present study were drawn from two different areas. There were students who answered the problems individually and those who solved the problems in groups. Data for the individual problem solving came from data gathered in 1976 by the Cognitive Analysis Project, CAP (Renner, 1977). Group problem solving data were collected in the spring semester, 1978.

The Aquarium and Hills problems were two science problems used in this research. Individual written solutions to these problems were obtained by the CAP in 1976 from 684 tenth, eleventh and twelfth grade students in the state of Oklahoma. These two science problems can be found in Appendices A and B. There were 366 students who answered the Aquarium problem and 318 students responded to the Hills problem. The criteria set by the CAP for these two problems were that the problems require a reading level not higher than seventh grade and that the science content be found in ninth and tenth grade science courses of study and textbooks. The solutions to these two problems provided by students during 1976 form the individual problem solving data for this study.

During the spring semester, 1978, students in five schools in the state of Oklahoma gave group responses to the Aquarium and Hills problems. These five schools were not used by the CAP. The schools were Shawnee High School (eleventh and twelfth grades), Central Innovative High School* in Oklahoma City (tenth, eleventh and twelfth grades), West Mid-High School in Norman (tenth grade biology), Moore High School (eleventh and twelfth grade biology), and Moore Central Mid-High School (tenth grade biology). The Principal of each school was asked to supply a random sample of students. There is no evidence to suggest that this request was not fulfilled.

---

*Central Innovative is a high school developed to provide alternative forms of education to the traditional high schools. A description of this school's program is given in Appendix C of the dissertation.

The biology classes at West Mid-High School (Norman) were selected because their students were experiencing a biology class which is oriented entirely to group work. Students at Moore Central Mid-High School were experiencing a traditional* biology class in which students work individually. Students from these two schools were used to investigate a possible relationship between previous group experience and ability to solve the problem. Furthermore, the group size of four to be used throughout the study was selected because this size was used in the West Mid-High School biology classes. Any groups formed with more or less than four students from different grade levels were excluded from the data. Students at the other schools do not have courses which emphasize group work as done at West Mid-High School. A description of the group work at West Mid-High School is given in Appendix D.

Piagetian and CAP Measures of Intellectual Development

Because this study was concerned with how the level of intellectual development might influence groups and individuals solving problems, measures of each participating

---

*In the traditional biology class the teacher explained biological concepts which were described in the selected text. Students in the class were treated as one large group. Most of the class period was spent by students listening to the teacher's exposition of the course content. The course content was primarily presented to the students through the teacher's directions and descriptions.

student's level of intellectual development had to be obtained.

Data for the students who solved the two science problems individually were collected in 1976 by the CAP. In addition to the students' responses to the problems, these data included a measure of each student's level of intellectual development from their performance on four Piagetian task interviews. Scores on these tasks ranged from four to fifteen. This range was divided by the CAP into three levels of intellectual development as follows:

4-8 : Concrete operational reasoning.

9-11: Transitional reasoning.

12-15: Formal operational reasoning.

During spring, 1978, the data collected from the groups of students included the responses of the groups to the problems as well as a measure of each student's level of intellectual development. Instead of using the Piagetian task interviews, each student who solved the problem in a group had his/her intellectual development obtained by a measure produced by the CAP. This CAP measure has three separate parts--two problems dealing with proportional reasoning (shadows (S), frogs (F)), one dealing with separation and control of variables (geraniums (G)) and the group embedded figures test (EFT) which measures the field independence* of a student. The regression equation which

*The embedded figures test is a perceptual test which requires the student to locate a simple figure within an predicts the level of intellectual development (EI score) is given by the following equation (Renner, 1977, p. 124):

$$EI = 0.17\ (EFT) + 0.38\ (S) + 0.37\ (F) + 0.30\ (G) + 3.95$$

The different levels of intellectual development according to the CAP measure have the same range of score as those mentioned above for the four Piagetian task interviews. This measure was obtained prior to the group solving the problems.

It was necessary to use two different measures of intellectual development even though this could be a confounding variable in the experiment. Piagetian task interviews require interviewing every student individually for approximately thirty minutes. Trained personnel and special equipment are required to administer these tests. Originally it was proposed to obtain data from approximately 125 groups of students. If 125 groups of four students were to be part of the study, 500 students would have to have been individually interviewed. Resources necessary for such large scale interviewing were unavailable. The advantage of the CAP measure is that it can be administered to a large group of students in fifty minutes. These tests were scored

---

obscuring, larger, complex figure. The student has to disembed the simple figure. Field dependent students focus on the overall organization of the field surrounding the simple figure, whereas field independent subjects can, by breaking up the organization of the field, locate the simple figure (Witkin et al., 1971).

by the investigator and also independently by a person who helped develop the scoring system for the CAP. Everyone involved with scoring the CAP measure in this research had been a part of the development of the CAP. Any scoring discrepancies between the raters were resolved by discussion.

Because two measures of intellectual development were used in this research it was necessary to determine the relationship between these measures. The problem was to determine whether a student who was classified as concrete, transitional or formal by the Piagetian task interviews received the same classification from the CAP test. To resolve this problem, the two measures were required on the same student. During 1976 the CAP did, in fact, collect these data on 142 tenth, eleventh and twelfth grade students. These students, however, were not the same students used in the individual problem solving section of this study. Three different statistical analyses were carried out on the data collected from the 142 students.

During 1976 the CAP found a correlation coefficient of 0.7 (Renner, 1977, p. 124) between the Piagetian task interviews and the CAP measure. This correlation coefficient was calculated across all three levels of intellectual development. Separate Pearson correlation coefficients between the two measures for each level of intellectual development had not been calculated by the CAP. For the purpose of the present study, these coefficients were calculated to determine the linear relationship between each level of intellectual development. A computer program written to compute these coefficients can be found in Appendix H. The Pearson correlation coefficient r is (McCall, 1975, p. 119):

$$r = \frac{n\Sigma X_i Y_i - (\Sigma X_i)(\Sigma Y_i)}{\sqrt{[n\Sigma X_i^2 - (\Sigma X_i)^2][n\Sigma Y_i^2 - (\Sigma Y_i)^2]}}$$

where: n = number of pairs of scores.

X_i = i^{th} student's score on one variable.

Y_i = i^{th} student's score on a second variable.

A χ^2 test of independence was also carried out to test the null hypothesis that the measures are independent. The χ^2 statistic is (McCall, 1975, p. 303):

$$\chi^2 = \sum_{k=1}^{C} \sum_{j=1}^{R} \frac{(O_{jk} - E_{jk})^2}{E_{jk}} \; d \; \chi^2_{(R-1)(C-1)}$$

where: O_{jk} = the observed frequency in the cell corresponding to the intersection of the j^{th} row and k^{th} column.

E_{jk} = the expected frequency in the cell corresponding to the intersection of the j^{th} row and k^{th} column.

R = the number of rows.

C = the number of columns.

Finally, the probability of a student scored at one level by one measure obtaining the same score from the other measure was calculated.

Formation of the Groups for the Problem Solving Session

During the spring semester, 1978, each school was visited twice--first to measure each student's level of intellectual development and second to obtain group solutions to the two problems. Some students who attended the first session were absent from the second session. Between the two visits to the schools the level of intellectual development of each student was calculated. A computer program written to compute the intellectual levels and provide a frequency distribution can be found in Appendix H. The types of groups according to degree of intellectual heterogeneity which might be formed was determined. At the beginning of the second session any absentees were noted. The remaining students were then assigned to the different groups as they entered the room. The allocation of a student to a particular group was not predetermined.

A question investigated in this study was whether it is better to have students who are at the same level of intellectual development solving problems in a group or whether the group should comprise students of varying levels of intellectual development. This problem was investigated by measuring the individual student's level of intellectual development and then forming groups in which students are heterogeneously or homogeneously grouped according to their level of intellectual development. There are three ways of homogeneously grouping students. All students in the group can be either at the concrete, transitional or formal levels of intellectual development. Likewise, there are four ways of forming intellectually heterogeneous groups. These four ways are:

(i) two concrete students with one transitional student and one formal student (heterogeneous-concrete)

(ii) two transitional students with one concrete student and one formal student (heterogeneous-transitional)

(iii) two formal students with one concrete student and one transitional student (heterogeneous-formal)

(iv) any combination of concrete, transitional and formal students (heterogeneous-miscellaneous).

In an effort to have equal numbers of each different intellectual type throughout each grade, a running tally of each of the different types was kept. Prior to each school visit the number and types of intellectual groups necessary to have an even distribution throughout the grades was determined. The groups at the school were compiled as close as possible to meet this condition.

Problem Solving Hierarchies

Hierarchies of responses which reflected different qualities of solutions to the problems were developed for the Aquarium and Hills problems. These hierarchies may be found in Appendices E and F. The procedure for developing these hierarchies was the same as used by the CAP. Four people who developed the CAP scoring categories for the tests repeated the procedure for this study. Initially all the individual and group responses to one of the problems were combined. A random sample of approximately 100 responses was read by four people, hereafter referred to as evaluators, and the responses divided into mutually exclusive divisions reflecting the different types of responses. A hierarchy of responses was developed to accommodate every response in the sample. A second random sample of approximately 100 responses was then drawn and the responses scored using the first hierarchy. If necessary the hierarchy was modified to accommodate every response in both samples. This procedure was repeated until every response could be consistently classified. The consistency achieved when a hierarchy was completed was at least ninety percent; that is, ninety percent or more of the evaluators gave the same score to a given response. The evaluators preparing the hierarchies were not concerned with instructing other persons outside the group how to use the hierarchies to score the solutions. The major concern was that every member of the group understand the categories and that internal consistency be achieved in the scoring procedure. A procedure to obtain categories worded such that people outside the group could quickly and easily score the solutions would have required several drafts and consultations. It was believed that if the study were to be replicated in the future other problems, and hence new hierarchies, would have to be developed. Nevertheless examples of student responses and scores have been given in Appendices E and F.

Variables in the Study

The present research contained a number of separate experiments each with its own set of hypotheses and appropriate experimental design. There was a number of independent variables in the study which had to be controlled either by randomization, blocking or by eliminating a particular group of students with a given set of characteristics from the analysis. These independent variables were:

(1) students solving problems individually or in groups of four.

(2) the student's grade--tenth, eleventh and twelfth.

(3) the student's level of intellectual development--concrete, transitional or formal.

(4) the degree of intellectual heterogeneity within the group or the intellectual composition of the groups.

Students were placed in intellectually homogeneous or intellectually heterogeneous groups. There were three different types of homogeneous groupings--concrete, transitional and formal. There were also four different heterogeneous groups--concrete, transitional, formal and miscellaneous. The composition of each of these types of intellectual groupings was described in the first section of this chapter.

The two dependent variables in this study were the scores received in the Aquarium or the Hills problems. Scores awarded to students who solved the problems individually became the data for the individual problem solving performances. Each group's score was the dependent variable for the group data. A group score had to be used because the statistical analyses required that the dependent variables be made up of independent or unrelated observations. If each student in a group received the score awarded to the whole group then each score or observation would not be independent. This is because all the members of the groups combined their efforts to formulate the group response. Each group, however, responded to the problems independently, without collusion. Therefore, the score on each problem received by the group became the observation or unit of analysis for the dependent variable which dealt with the group data.

Hypotheses

Alternative hypotheses were stated so that the anticipated direction of the difference between the means may be evident. The statistical tests were, however, carried out using the null hypothesis that individual and group performances were in fact equal (the equality point). If the null hypothesis is rejected at the equality point the hypothesis can be rejected for any other point that is in the opposite direction from the alternative hypothesis prediction.

Therefore, the hypothesis that the group performance is less than the individual performance could be rejected.

1. H_1: For students without previous group experience the mean score on the Aquarium Problem obtained by students who solve the problem in groups will be higher than the mean score obtained by students who solve the problem individually.
2. H_1: Students without previous group experience at higher levels of intellectual development as measured by Piagetian task interviews or the CAP measure of intellectual development will obtain a higher mean score on the Aquarium Problem than students at lower levels of intellectual development.
3. H_1: Students without previous group experience in higher grades will obtain a higher mean score on the Aquarium Problem than students in the lower grades.
4. H_1: The mean score on the Aquarium Problem obtained by students who are intellectually homogeneously grouped will be less than that obtained by students who are intellectually heterogeneously grouped. These students do not have previous group experience. Levels of intellectual development will be measured by tasks derived from Piagetian theory.

5. H_1: The mean score on the Aquarium Problem obtained by tenth grade students who are experiencing a biology class based on group experience will be higher than the mean score obtained by tenth grade students who are experiencing biology which is not based on group work.

6-10. H_1: The alternative hypotheses 6-10 will be the same as hypotheses 1-5 except that they will investigate student performance on the Hills Problem.

The Experimental Designs for Testing the Hypotheses

Computer programs, written to obtain the random samples for testing these hypotheses, may be found in Appendix H.

Hypotheses 1-3 and 6-8

Answers to the following questions were sought for students without previous group experience and intellectually homogeneous groups:

(i) Do students who answer the Aquarium problem in a group receive higher scores than students who answer the same problems individually? (Hypothesis 1)

(ii) Do students at higher levels of intellectual development receive higher scores on the Aquarium problem than students at lower intellectual levels? (Hypothesis 2)

(iii) Do students in higher grades receive higher scores on the Aquarium problem than students in lower grades? (Hypothesis 3)

Questions similar to (i), (ii) and (iii) above were investigated but they dealt with the Hills problem rather than the Aquarium problem. (Hypotheses 6, 7 and 8)

It was noted earlier that the West Mid-High School students were the only students in this study who had been deliberately taught in groups and have had considerable group experience. As far as is known, the students in the CAP who answered the problems individually had not had planned group experience similar to the West Mid-High School students. Therefore, West Mid-High School students were excluded from the analysis in this section of the study.

Level of intellectual development was one of the independent variables investigated in this study. When this variable was studied only intellectually homogeneous groups were used for the group problem solving data. Each student who formed part of the individual problem solving data was categorized as at the concrete, transitional or formal operational level of intellectual development. Not all of the groups, however, could be categorized in a manner equivalent to the students who solved the problems individually. Intellectually heterogeneous groups as defined in the first section of this chapter were composed of students at more than one level of intellectual development. For example, heterogeneous-concrete groups contained two concrete operational students, one transitional and one formal operational student. Such a group is neither equivalent intellectually to a homogeneous-concrete group which contains concrete operational students nor to an individual concrete operational student. Therefore only homogeneous groups could be equated intellectually to the three levels of mental development into which the students who answered the problems individually were classified. The group problem solving data for testing hypotheses one to three and six to eight were collected from intellectually homogeneous groups only.

A three-way ANOVA was planned to test hypotheses 1-3 and 6-8. The independent variables were groups without previous group experience against individuals without previous group experience, level of intellectual development and grade (ten, eleven and twelve).

Hypotheses 4 and 9

Do intellectually homogeneous groups without previous group experience score higher on the Aquarium or Hills Problem than intellectually heterogeneous groups without previous group experience? This was the question posed by hypotheses 4 and 9. To answer this question all of the group responses from each grade were categorized into intellectually homogeneous and heterogeneous groups. Then a two-way ANOVA with the independent variables degree of intellectual heterogeneity in the group and grade was to be performed.

Again, data from West Mid-High School students were excluded because of their previous group experience. Control of the independent variable--grade--was achieved by using it as a blocking variable.* Renner et al. (1976, p. 102) demonstrated that level of intellectual development varies with grade. That problem solving success is related to level of intellectual development was shown by the CAP (Renner, 1977). Therefore, possibly grade level may be positively correlated with ability to solve problems used in this research. Ideally level of intellectual development should also be used as a blocking variable but, as explained earlier, the four different types of intellectually heterogeneous groups cannot be equated with the three different types of intellectually homogeneous groups.

Hypotheses 5 and 10

Do tenth grade biology students who are experiencing group work score higher on the Aquarium or Hills Problem than tenth grade biology students who are not experiencing group work? An answer to this question was sought by testing hypotheses five and ten with a one-tailed t-test. A computer program, written to perform this test, may be found in Appendix H.

Data were selected from the West Mid-High School and Moore Central Mid-High School tenth grade biology classes.

---

*In research the variability among subjects may obscure the treatment effects being investigated. If this variability is caused by a nuisance variable which can be identified the variability among subjects can be minimized. Subjects are assigned to blocks or categories of this nuisance or blocking variable so that the variability among subjects within any block is less than the variability among the blocks. This blocking variability is thought to be highly correlated with the dependent variable.

Each group of students was classified according to the type of previous group experience. The independent variable, degree of intellectual heterogeneity within the group, was to be controlled by randomization.

Statistical Considerations

Type I and Type II Errors

Before establishing the level of significance to be used to test the hypotheses, consideration was given to the types of error which may occur. There are two types of error--Type I and Type II. A Type I error occurs when the null hypothesis is falsely rejected. On the other hand, if the null hypothesis is falsely accepted, a Type II error occurs.

In this research if the null hypothesis that there are no differences between groups and individuals was falsely rejected, a Type I error would have occurred. In other words, the erroneous conclusion would be drawn that there are differences between groups and individuals solving problems. On the basis of this conclusion teachers would change their method of instruction when, in fact, there was no need, because it makes no difference to the students how they are taught. Hence, if a Type I error was made, students have not been harmed educationally when the erroneous conclusion is implemented in the classroom.

If the null hypothesis was actually false and a Type II error was committed by accepting this null hypothesis, the conclusion of the study would be that there were no differences between groups and individuals solving problems. If this conclusion were implemented in schools, teachers would not change their teaching methodology when in fact they should. Hence, the students would not benefit from the results of the research and would therefore be educationally deprived. Clearly a Type II error should be avoided when drawing conclusions from the results of testing the hypotheses one and six which test group and individual performance at solving problems.

Power and Cell Sizes

Type II errors cannot be directly controlled by the investigator. Nevertheless, the probability of a Type II error is related to the level of significance and the power of the statistical test. A discussion of this relationship can be found in Appendix K. The higher the power of a statistical test, the lower the probability of a Type II error. It may therefore seem desirable for the tests to be used in this research to have a power of 0.999. This would result in virtually any differences between the means of the populations being detected as statistically significant. Such minute differences may well be meaningless and unrelated to the effects of the independent variable on the dependent variable. Instead, a power of 95 per cent at a level of significance of 0.05 was selected for all the statistical tests in this research. Cell sizes required to detect

1.0 σ_e* and 1.5 σ_e differences were computed. The tests were to be carried out to detect differences of 1 σ_e, but if this were not possible the smaller cell sizes required for differences of 1.5 σ_e were to be used. These cell sizes are shown in Table 1. See Appendix J for the computational formulae.

TABLE 1

COMPUTED CELL SIZES FOR TESTING THE HYPOTHESES

σ_j	Hypotheses				
	1,6	2,7	3,8	4,9	5,10
1.0 σ_e	3	3	3	9	27
1.5 σ_e	1	1	1	4	13

where σ_j = j^{th} treatment effect

σ_e = population error standard deviation

It was anticipated that there could be insufficient data to carry out the three-way ANOVA to test hypotheses one to three and six to eight to detect a difference of 1.0 σ_e. The alternative of carrying out the test to detect a difference of 1.5 σ_e was unacceptable because a random sample of one was not regarded as representative of the population. Another possibility was to test hypotheses one, two, six and seven by a two-way ANOVA. For this test cell sizes of three

---

* σ_e = the population error standard deviation. The relationship between σ_e, power and cell size is discussed in Appendix K.

would be required to have 85 per cent power at $\alpha = .05$ and a difference of 1.5 σ_e. The power of this test could be increased by selecting cell sizes for the individual factor proportionally greater than the cell sizes for the group factor. A proportional ANOVA could then be used to analyze the data. Individual cell sizes of nine would increase the power of the test because these cell sizes for a two-way ANOVA would result in a test having a power of 95 per cent at $\alpha = 0.05$ and a difference of 1.0 σ_e. If necessary a t-test could compare group and individual performances. To achieve 95 per cent power at $\alpha = 0.05$ cell sizes of twenty-seven and thirteen would be required to detect differences of 1 σ_e and 1.5 σ_e, respectively.

Because of the uncertainty of obtaining the desirable cell sizes as planned, random samples were selected from the individual data so that four different tests could be carried out. Their tests were the three-way ANOVA with cell sizes of three, the two-way ANOVA with cell sizes of nine and two t-tests with cell sizes of thirteen and twenty-seven. A computer program, found in Appendix H, was prepared to obtain these random samples. All of these responses were then scored using the hierarchies.

Meeting the Assumptions of the Statistical Tests

The experimental procedures discussed in the fourth section were planned so that the assumptions of the statistical tests described in that section would not be violated so that the tests were invalid. The assumption of independence of observations was obtained with the group data by using the group score as the unit of analysis. Scores obtained from students who responded individually to the problems were independent because there was no collusion amongst the students. Violation of the assumption of homogeneity of variance was to be avoided by randomly selecting observations to obtain equal cell sizes. A more detailed discussion of these underlying assumptions is given in Hays (1973, pp. 481-483).

References

Hays, W. Statistics for the Social Sciences, 2nd ed. New York: Holt, Rinehart and Winston, Inc., 1973.

McCall, R. B. Fundamental Statistics for Psychology, 2nd ed. New York: Harcourt, Brace Jovanovich, Inc., 1975.

Pearson, E. S. and Hartley, H. O. Charts of the Power Function for the Analysis of Variance Tests Derived from the Non-Central F Distribution. Biometrika, 38, 1951.

Renner, J. W. Evaluating Intellectual Development Using Written Responses to Selected Science Problems, The Story of the Cognitive Analysis Project. Norman: The University of Oklahoma, 1977.

Renner, J. W., et al. Research, Teaching and Learning with the Piaget Model. Norman: The University of Oklahoma Press, 1976.

Witkin, H. A., Oltman, P. K., Raskin, E. and Kays, S. A. A Manual for the Embedded Figures Tests. Palo Alto, California: Consulting Psychologists Press, Inc., 1971.

CHAPTER IV

RESULTS AND INTERPRETATION OF THE RESEARCH

The results of the present study will be divided into four major sections. First, the two measures of intellectual development--the Piagetian task interviews and the CAP measure--will be compared. Secondly, the hierarchies representing different qualities of responses to the two science problems will be discussed. Thirdly, the individual and group data will be presented. Finally, the hypothesis testing will be presented and the results interpreted.

A Comparison of the Piagetian and the CAP Measures of Intellectual Development

There were two measures of the level of intellectual development, the Piagetian task interviews and the CAP measure, available for each of 142 students who participated in part of the 1976 CAP. Statistical tests were carried out to answer the following questions:

What is the relationship between performance on the Piagetian task interviews and the CAP measure?

What is the probability of a student who is rated at one particular level of intellectual development by the

Piagetian measure receiving the same categorization by the CAP measure?

To answer the first question two statistical tests were carried out. These were the calculation of the Pearson correlation coefficient and the test of independence using the χ^2 statistic. Separate correlation coefficients were calculated for the whole sample across all levels of intellectual development and then for each separate level of intellectual development. The results are shown in Table 2. Data from which these results were obtained are given in Appendix G.

TABLE 2

CORRELATION COEFFICIENTS FOR THE PIAGETIAN TASKS AND THE CAP MEASURES

	All 3 Levels	Concrete	Transitional	Formal
n	142	48	61	33
r	0.7	0.3	0.4	0.5

The correlation coefficient obtained for the whole sample across all three levels of intellectual development agrees with the result obtained by the CAP. Lower correlations for each separate level of intellectual development were expected because only part of the total scale for the measure of intellectual development was used in the analysis. Restriction in the range of the scale tends to lower the correlation coefficient. Nevertheless, all correlation coefficients were positive. This indicated that there was a positive, linear relationship between the two measures of intellectual development.

The chi-squared (χ^2) test investigated the degree of independence of the two measures. The null hypothesis tested was that a student's score on the Piagetian tasks was independent of that same student's score on the CAP test. Results of the χ^2 test are shown in Table 3.

TABLE 3

FREQUENCY TABLE OF PIAGETIAN AND CAP LEVELS

CAP Level	Piagetian Level			Total
	Concrete	Transitional	Formal	
Concrete	32	13	2	47
Transitional	16	43	18	77
Formal	0	5	13	18
Total	48	61	33	142

$\chi^2_{obs} = 58.67$ Critical value $\chi^2_{\substack{\alpha=.05 \\ df=4}} = 9.49$

The null hypothesis was rejected. Therefore, from the results of the two statistical analyses, it was concluded that the two measures of intellectual development were related.

To answer the second question a number of probabilities was calculated. For example, the probability was calculated of a student who was categorized as concrete by the

Piagetian task interview receiving the same categorization according to the CAP measure. From Table 3 it can be seen that out of a total of forty-eight students who were classified at the concrete level of intellectual development on the Piagetian measure, thirty-two of these students received the same classification on the CAP measure. The probability of this event occurring is 32/48 = 0.67. Similar probabilities were calculated for the other categories. These results are shown in Table 4.

TABLE 4

PROBABILITIES OF STUDENTS AT PARTICULAR PIAGETIAN LEVELS RECEIVING A PARTICULAR CAP CLASSIFICATION

Piaget Classification	CAP Classification		
	Concrete	Transitional	Formal
Concrete	0.67	0.33	0.00
Transitional	0.21	0.71	0.08
Formal	0.06	0.55	0.39

From Table 4 it can be seen that the two measures are comparable for students at the concrete and transitional stages according to the Piagetian task interviews. The probabilities that these students would receive the same classification in the CAP and Piagetian measures were 0.67 and 0.71, respectively.

There was a probability of 0.21 that a student categorized as at the transitional level by the Piagetian task interviews would be scored concrete by the CAP measure. It would appear then that the CAP measure tends to give a lower categorization for students at the transitional level according to the Piagetian task interviews. One explanation for this could be the different natures of the two measures. The Piagetian task interviews involved a personal interview with the student. During this time rapport between the student and interviewer could be established. The nature of and reasons for the tasks could be explained to the student. This situation might encourage or motivate the student. On the other hand, the CAP measure was administered to a large group of students during which time each student provided individual written solutions to the problems. The administrative advantage of being able to measure the intellectual development of a large number of students at the one time may have the disadvantage of removing the motivational aspects of a personal interview. Furthermore, the Piagetian task interviews enable the interviewer to structure the questions according to the student's responses. The interviewer is free to probe the student's understanding of the tasks. This freedom and flexibility is not available in a mass-testing situation. There is also inherent in the CAP measure itself an error of measurement. The CAP reported a standard error of measurement of the CAP test of 1.85 (Renner, 1977, p. 124). A similar error of measurement is not available for the Piagetian task interviews. It is a premise of this research that these task interviews are valid measurements of the levels of intellectual development.

Refer again to Table 4. The probability of a student categorized at the concrete stage by the Piagetian task interviews being classified as transitional by the CAP test was 0.33. In this instance the CAP measure was overscoring relative to the Piagetian measure. As mentioned earlier there is an error of measurement in the CAP measure. Therefore, because there is no lower category than concrete on the CAP scale, any error in the measurement will be evident by overscoring at the concrete level. It should be noted that the error in measurement resulted in an overscoring of just one level, from concrete to transitional, not of two levels, concrete to formal.

Students who were classified at the formal level of intellectual development by the Piagetian task interviews had a probability of 0.55 of being classified as transitional by the CAP measure. This underscoring of the CAP measure may have resulted from several causes. First of all, the formal category is the highest level of intellectual development in both measures. Hence, any error in the CAP measure will be evident as underscoring. Furthermore, as discussed earlier, the CAP measure may underscore because of the loss of the motivational aspect of a personal interview and the constraints of a testing situation which does not allow an interviewer to probe a student's understanding of the problem. As with the concrete stage, the misclassification of the formal students by the CAP test was mainly just for one level (formal to transitional) not two levels (formal to concrete).

Generally the CAP measure and the Piagetian task interviews are dependent and give nearly the same measures of intellectual development, except for a tendency to rate students classified as formal by the Piagetian task interviews as transitional on the CAP measure. The underscoring by the CAP measure at the formal level may not effect a large percentage of the students. By using the Piagetian task interviews, Renner et al. (1976, pp. 94-95) determined the levels of intellectual development of high school students in the state of Oklahoma. The percentage of students at each level of intellectual development according to grade is given in Table 5.

TABLE 5

PERCENTAGE OF STUDENTS AT DIFFERENT LEVELS OF INTELLECTUAL DEVELOPMENT ACCORDING TO GRADE, AS MEASURED BY PIAGETIAN TASK INTERVIEWS

Piagetian Task Classification	Grade		
	10	11	12
Concrete	73.40	68.69	65.98
Transitional	15.96	19.19	15.46
Formal	10.64	12.12	18.56

There was a small percentage of students at the formal level, even at the twelfth grade. Hence, any misclassification at the formal level by the CAP would only affect a small percentage of the students.

The Scoring Hierarchies

Separate hierarchies both with five categories were developed for the science problems. These two hierarchies are given in Appendices E and F. The categories in each hierarchy were sequenced to represent an increase in the quality of the solution to the problem. Responses in the lowest category were scored zero. They did not represent any thought related to the problem. It was felt by the group developing the hierarchies that points should be awarded only for relevant statements which represent different qualities of thought.

There were two types of group responses which could not be scored by the evaluators. Responses of these types were omitted from the study. First, writing on the answer sheet included two or more different types of handwriting and these responses demonstrated completely different levels of understanding of the problem. This suggested that two or more individuals had responded separately to the problem; i.e., there was no group response. Any data collected from this group was omitted from any further analysis. Separate scores could not be awarded to each written response, because there was no evidence to suggest which response was given by a particular student. Furthermore, the scores or observations of the dependent variable have to be independent or unrelated. This was necessary for the validity of the statistical tests. (See Chapter III, fifth section, "Statistical Considerations.") The second reason for omitting group responses from the data was that the group did not contain four students or that all the members of the group were not from the same grade.

The Aquarium Problem Hierarchy

To gain any points, the student had to present a response which indicated that the content of the problem had been recognized. If only irrelevant statements were made, no points were awarded (Category 0). One point was awarded to responses which discussed the problem. In this category--Category 1--the results of the experiment, the appearance of the green color in the aquarium, was explained. Experimental tests to find out why the green color occurred were not described. The next category of responses--Category 2--in the hierarchy included procedures which could partially solve the problem, but an organized approach to the problem was lacking. Responses in the next category--Category 3--correctly combined the different elements of the problem, but only two elements were combined at a time. In the last category--Category 4--the students, having considered the combination of two elements, extended their solutions to combine three or more elements.

The Hills Problem Hierarchy

Responses in the lowest category received no points--Category 0. To receive a point an individual or a group generally restated the problem or offered explanations for the difference in the ruts. Experimental proof was not part of the answer. These responses were in Category 1. In the next category--Category 2--the responses indicated some experimental manipulation of the variables. There was an organized attempt to solve the problem by setting up an experiment. The experiment, however, did not indicate a necessity to control the variables in the experiment. Such a solution would fail to solve the problem. Setting up models of the hills in the laboratory with some measure of control was a characteristic of responses in the next category--Category 3. These solutions were inadequate because of insufficiently controlled variables. The necessity for obtaining experimental measurements was recognized in this category. In the final category--Category 4--a valid experiment which would test the hypothesis was described. This response category also indicated the purpose of the experiment and what results would be expected if the hypothesis were to be substantiated.

The Group and Individual Data

Group Data

The number of the different types of groups used in the study is shown in Table 6. These figures represent the total group population available for the analysis of both problems.

TABLE 6

NUMBER OF DIFFERENT TYPES OF GROUPS AVAILABLE

Group Type	Grade				
	10			11	12
	School #			School #	School #
	2-5	1	5	2-4	2-4
Homo-C	4	2	3	10	2
Homo-T	4	3	4	7	3
Homo-F	1	0	1	1	1
Hetero-C	0	0	0	4	1
Hetero-T	0	7	0	2	5
Hetero-F	0	0	0	1	1
Hetero-M	2	2	0	6	5

Where: School #1 = West Mid-High School
School #2 = Moore High School
School #3 = Shawnee High School
School #4 = Central Innovative High School
School #5 = Moore Central Mid-High School

Group Type:

Homo-C = homogeneous-concrete

Homo-T = homogeneous-transitional

Homo-F = homogeneous-formal

Hetero-C = heterogeneous-concrete

Hetero-T = heterogeneous-transitional

Hetero-F = heterogeneous-formal

Hetero-M = heterogeneous-miscellaneous.

The definitions of each of these types of groups is given in Chapter III.

In addition to the scores received by each group for their solutions to the two problems, the school number, grade, group type and previous group experience are shown in Table 7.

TABLE 7

GROUP DATA FOR THE AQUARIUM AND HILLS PROBLEMS

Group #	School #	Grade	Group Type	Group Experience	Problem Score Aquarium	Problem Score Hills
1	1	10	1	1	2	2
2	1	10	1	1	4	4
3	5	10	1	0	1	2
4	5	10	1	0	1	2
5	5	10	1	0	1	2
6	4	10	1	0	2	2
7	1	10	2	1	2	2
8	1	10	2	1	3	3
9	1	10	2	1	4	1
10	5	10	2	0	4	2
11	5	10	2	0	4	3
12	5	10	2	0	0	1
13	5	10	2	0	3	2
14	5	10	3	0	2	3
15	1	10	5	1	3	3
16	1	10	5	1	4	0
17	1	10	5	1	3	4
18	1	10	5	1	3	2
19	1	10	5	1	3	3
20	1	10	5	1	4	4
21	1	10	5	1	4	3
22	1	10	7	1	3	3

TABLE 7--Continued

Group #	School #	Grade	Group Type	Group Experience	Problem Score Aquarium	Problem Score Hills
23	1	10	7	1	1	4
24	4	10	7	0	4	2
25	4	10	7	0	2	2
26	3	11	1	0	2	1
27	3	11	1	0	2	1
28	3	11	1	0	2	2
29	2	11	1	0	1	2
30	2	11	1	0	2	1
31	4	11	1	0	2	1
32	3	11	1	0	3	2
33	3	11	1	0	1	2
34	3	11	1	0	2	1
35	3	11	1	0	1	2
36	3	11	2	0	3	1
37	2	11	2	0	2	2
38	2	11	2	0	1	2
39	2	11	2	0	3	1
40	3	11	2	0	1	1
41	3	11	2	0	3	1
42	3	11	2	0	1	3
43	2	11	3	0	4	4
44	3	11	4	0	3	3
45	3	11	4	0	1	3

TABLE 7--Continued

Group #	School #	Grade	Group Type	Group Experience	Problem Score Aquarium	Problem Score Hills
46	3	11	4	0	3	2
47	4	11	4	0	2	1
48	3	11	5	0	2	0
49	2	11	5	0	4	3
50	3	11	6	0	3	2
51	3	11	7	0	1	1
52	3	11	7	0	1	1
53	4	11	7	0	4	0
54	4	11	7	0	1	2
55	2	11	7	0	1	2
56	2	11	7	0	3	2
57	3	12	1	0	2	1
58	3	12	1	0	2	2
59	3	12	2	0	4	1
60	2	12	2	0	2	3
61	2	12	2	0	3	4
62	2	12	3	0	2	3
63	3	12	4	0	1	2
64	2	12	5	0	4	2
65	2	12	5	0	4	2
66	3	12	5	0	4	3
67	3	12	5	0	3	0
68	3	12	5	0	2	0

TABLE 7--Continued

Group #	School #	Grade	Group Type	Group Experience	Problem Score Aquarium	Problem Score Hills
69	2	12	6	0	4	2
70	3	12	7	0	1	2
71	3	12	7	0	2	3
72	4	12	7	0	2	2
73	4	12	7	0	3	3
74	4	12	7	0	3	2

School #1 = West Mid-High School
School #2 = Moore High School
School #3 = Shawnee High School
School #4 = Central Innovative High School
School #5 = Moore Central Mid-High School

Group Type:

1 = Homogeneous-concrete
2 = Homogeneous-transitional
3 = Homogeneous-formal
4 = Heterogeneous-concrete
5 = Heterogeneous-transitional
6 = Heterogeneous-formal
7 = Heterogeneous-miscellaneous

Group Experience:

0 = no previous group experience
1 = has previous group experience

Individual Data

Random samples were selected from the 366 and 318 students who responded individually to the Aquarium and Hills problems, respectively. Tables 8 and 9 show the scores received by these individual students for the Aquarium and Hills problems, respectively, together with the student's grade and level of intellectual development as measured by the Piagetian task interviews.

TABLE 8

INDIVIDUAL DATA FOR THE AQUARIUM PROBLEM

Student #	Grade	Intel-lectual Level	Problem Score	Student #	Grade	Intel-lectual Level	Problem Score
1	10	5	0	23	11	8	2
2	10	6	1	24	11	8	3
3	10	8	2	25	11	5	1
4	10	5	1	26	11	5	1
5	10	8	2	27	11	8	0
6	10	8	2	28	11	6	1
7	10	7	2	29	11	8	4
8	10	7	1	30	11	7	4
9	10	7	1	31	11	6	1
10	10	5	0	32	11	10	1
11	10	8	2	33	11	11	1
12	10	7	1	34	11	10	1
13	10	6	1	35	11	10	1
14	10	10	3	36	11	11	2
15	10	11	2	37	11	10	2
16	10	9	1	38	11	9	3
17	10	9	2	39	11	10	1
18	10	11	3	40	11	11	2
19	10	9	1	41	11	12	2
20	10	9	1	42	11	15	3
21	10	10	2	43	11	12	3
22	10	12	3	44	11	13	3

TABLE 8--Continued

Student #	Grade	Intel-lectual Level	Problem Score	Student #	Grade	Intel-lectual Level	Problem Score
45	11	12	1	59	12	8	1
46	11	12	4	60	12	10	2
47	11	12	1	61	12	10	1
48	11	15	0	62	12	11	1
49	11	13	2	63	12	11	2
50	11	13	2	64	12	9	1
51	12	7	2	65	12	11	1
52	12	7	1	66	12	12	2
53	12	8	1	67	12	12	2
54	12	8	1	68	12	13	3
55	12	8	3	69	12	12	2
56	12	7	1	70	12	13	3
57	12	7	2	71	12	12	1
58	12	8	2	72	12	12	3

TABLE 9

INDIVIDUAL DATA FOR THE HILLS PROBLEM

Student #	Grade	Intel-lectual Level	Problem Score	Student #	Grade	Intel-lectual Level	Problem Score
1	10	7	1	23	11	7	1
2	10	8	2	24	11	8	3
3	10	7	1	25	11	8	0
4	10	7	1	26	11	8	1
5	10	8	3	27	11	6	1
6	10	8	0	28	11	5	1
7	10	11	1	29	11	5	1
8	10	9	2	30	11	8	2
9	10	10	1	31	11	6	1
10	10	11	2	32	11	8	3
11	10	9	1	33	11	10	2
12	10	10	2	34	11	9	1
13	10	9	2	35	11	11	2
14	10	9	1	36	11	9	1
15	10	11	1	37	11	10	2
16	10	10	1	38	11	9	1
17	10	10	1	39	11	11	1
18	10	13	2	40	11	9	4
19	10	15	2	41	11	9	2
20	10	14	2	42	11	13	2
21	10	8	2	43	11	13	1
22	11	5	2	44	11	12	1

TABLE 9--Continued

Student #	Grade	Intel-lectual Level	Problem Score	Student #	Grade	Intel-lectual Level	Problem Score
45	11	12	4	58	12	11	2
46	11	12	2	59	12	10	0
47	11	13	3	60	12	11	2
48	11	12	1	61	12	9	2
49	11	12	2	62	12	9	0
50	12	5	2	63	12	14	4
51	12	7	2	64	12	15	3
52	12	8	1	65	12	14	2
53	12	7	1	66	12	13	2
54	12	6	1	67	12	13	2
55	12	8	1	68	12	12	3
56	12	11	1	69	12	13	2
57	12	10	2	70	12	13	3
				71	12	12	1

Results of Testing the Hypotheses

This section discusses the results of testing the hypotheses one through ten and how the results of that testing lead to additional hypotheses being postulated and tested. A summary of the results of the hypothesis testing will be given.

The cell sizes necessary to test hypotheses one through ten with 95 per cent power and differences of 1 or 1.5 σ_e at $\alpha = 0.05$ were calculated prior to analyzing the data (Table 1). After collecting the data it was found that the cell sizes were insufficient to analyze the data as planned. Reasons for these insufficiencies are discussed in the second section of this chapter. Different tests were tnerefore carried out. These tests increased the cell sizes by either reducing the number of independent variables in the study or by using all of the appropriate data available for the tests. The results of the statistical analyses, together with the cell means, will be presented in this section. The data from which these results were obtained are given in Appendix G.

The powers of the statistical tests were calculated using the non-centrality parameter, ϕ and the Pearson-Hartley charts (Pearson and Hartley, 1951). These powers were calculated for $\alpha = 0.05$. If the tests were rejected at levels of significance greater than 0.05 the powers of the tests would be increased. The exact powers, however, could not be obtained, because the Pearson-Hartley charts are only available for α-values of 0.01 and 0.05.

In this chapter the level at which each test was significant will be reported. Frequently in research, decisions to reject the null hypotheses at particular significance levels are made either by following the convention of using α-values of 0.01 or 0.05 or by using the same values reported in previous research. It was believed that, for the present research, educational implications and the probability of avoiding Type II errors should be the criteria for deciding the level of significance at which to reject the null hypotheses. The decisions to reject the null hypotheses according to these criteria will be discussed in Chapter V in the section "Conclusions and Educational Implications."

Hypotheses 1-3, 6-8

The results of testing the following hypotheses will be discussed in this section.

H_1:1 For students without previous group experience the mean score on the Aquarium problem obtained by students who solve the problem in groups will be higher than the mean score obtained by students who solve the problem individually.

H_1:2 Students, without previous group experience, at higher levels of intellectual development as measured by Piagetian task interviews or the CAP measure of intellectual development will obtain a higher mean score on the Aquarium problem than students at lower levels of intellectual development.

H_1:3 Students without previous group experience in higher grades will obtain a higher mean score on the Aquarium problem than students in lower grades.

Hypotheses 6, 7 and 8 are similar to those above but deal with the Hills problem.

Originally, it was planned to test these hypotheses in a three-way ANOVA with cell sizes of three (Table 1). The independent variables were groups without previous group experience against individuals, without previous group experience, intellectual level and grade. It can be seen from Table 6 of this chapter that, if only homogeneous groups were used, the maximum cell size available for the homogeneous-formal groups was one. Because only insufficient data were available, the decision was made to collapse across grade levels and carry out a two-way ANOVA with the independent variables: groups without previous group experience versus individuals without previous group experience and levels of intellectual development. The reasons for this lack of data are explained in the second section of Chapter IV. This test would give a maximum cell size of three in the homogeneous-formal category. Rather than randomly selecting cell sizes of three for the homogeneous-concrete and homogeneous-transitional groups, the decision was made to use the maximum possible cell sizes for these two categories. These cell sizes would be sixteen and fourteen, respectively. Such a procedure would provide a more powerful test than one with all cell sizes of three. Random samples of the individual data were drawn from those individuals whose problem solutions had already been scored. Cell sizes of sixteen, fourteen and three were selected for each of the three categories of individuals in the concrete, transitional and formal stages of intellectual development, respectively. The data were then analyzed with a two-way ANOVA using proportional cell sizes. Equations for this analysis are given in Kirk (1968, pp. 200-202). Tables 10 and 11 give the cell means for each problem. The results of the analyses for both the Aquarium and Hills problems are given in Tables 12 and 13, respectively.

In eight of the cell means in Tables 10 and 11, the cell means for the group performances were higher than those of the individual performances. Therefore, irrespective of the levels of intellectual development, the groups solved the problems better than the individuals. There were two exceptions. Individuals at the concrete operational level performed better on the Aquarium problem than the homogeneous-concrete groups. Concrete operational individuals performed as well as homogeneous-concrete groups on the Hills problem.

It was expected that students at higher levels of intellectual development would perform better than students

TABLE 10

CELL MEANS FOR THE 2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND INTELLECTUAL LEVEL--AQUARIUM PROBLEM

Level of Intellectual Development	Groups	Individuals
Concrete	1.7	1.8
Transitional	2.4	1.5
Formal	2.7	2.3

TABLE 11

CELL MEANS FOR THE 2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND INTELLECTUAL LEVEL--HILLS PROBLEM

Level of Intellectual Development	Groups	Individuals
Concrete	1.6	1.6
Transitional	1.9	1.6
Formal	3.3	3.0

TABLE 12

SUMMARY TABLE: 2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND INTELLECTUAL LEVEL--AQUARIUM PROBLEM

Source	S.S.	d.f.	M.S.	F.
Group vs. Individual	2.56	1	2.56	2.71 (1)
Intellectual Level	3.32	2	1.66	1.76 (2)
Interaction	3.67	2	1.84	1.94 (3)
Error	56.7	60	0.95	
Total	66.25	65		

(1) $p<0.2$ (2) $p<0.2$ (3) $p<0.2$

TABLE 13

SUMMARY TABLE: 2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND INTELLECTUAL LEVEL--HILLS PROBLEM

Source	S.S.	d.f.	M.S.	F
Group vs. Individual	0.55	1	0.55	0.77 (1)
Intellectual Level	12.55	2	6.28	8.86 (2)
Interaction	0.22	2	0.11	0.16 (3)
Error	42.50	60	0.71	
Total	55.82	65		

(1) $p>0.25$ (2) $p<0.001$ (3) $p>0.25$

at lower levels of intellectual development. A comparison of the cell means in Tables 10 and 11 show that this relationship generally holds. For both individual and group performances, the formal operational students received higher scores than transitional students, who in turn received higher scores than concrete operational students. Again there were two exceptions. Concrete operational students who answered the Aquarium problem individually performed better than transitional students who answered the same problem individually. Furthermore, concrete operational individuals who answered the Hills problem received the same mean score as transitional students who answered the problem individually.

Consider the results of the data analysis in Tables 12 and 13. The power of these statistical tests for the group versus individual variable was 98 percent at $\alpha = 0.05$ for a standard deviation of 1 σ_e. The F-ratio for the group versus individual independent variable was significant at $\alpha = 0.20$ for the Aquarium problem (critical $F^{0.20}_{1,60} = 1.68$). Thus, if the null hypothesis for the Aquarium problem is rejected at $\alpha = 0.20$, this result means that groups without previous group experience perform significantly better on this problem than individuals without prior group experience. It would appear that the action of forming groups does in fact lead to superior performances on solving the Aquarium problem. Whether the null hypothesis should be rejected is discussed in Chapter V. The results for the Hills problem, however, lead to a different interpretation. The F-ratio for this problem of the group versus individual variable was not significant at $\alpha = 0.25$ (critical $F^{0.25}_{2,60} = 1.35$). Reasons for not rejecting the null hypothesis associated with the Hills problem are discussed in Chapter V. Therefore, for the Hills problem, there was no significant difference between the performances of groups and individuals, both without previous group experience. Even though the group had a higher mean, the difference between the groups and individuals was not statistically significant.

The results of investigating the effects of level of intellectual development on problem solving performance are reported in Tables 12 and 13. The F-ratios for the Aquarium and Hills problems are significant at α-values of 0.20 and 0.001 respectively (critical $F^{0.20}_{2,60} = 1.65$ and critical $F^{0.001}_{2,60} = 7.76$). It was decided to reject the null hypothesis for the Hills problem, but not for the Aquarium problem. These decisions regarding rejection of the null hypotheses are discussed in Chapter V. The results of these analyses indicate that the scores received by problems on the Hills problem are related to the students' levels of intellectual development. Furthermore, from Table 11 it appears that the higher a student's level of intellectual development, the higher the score received on the Hills problem by that student. These relationships did not occur with the Aquarium problem.

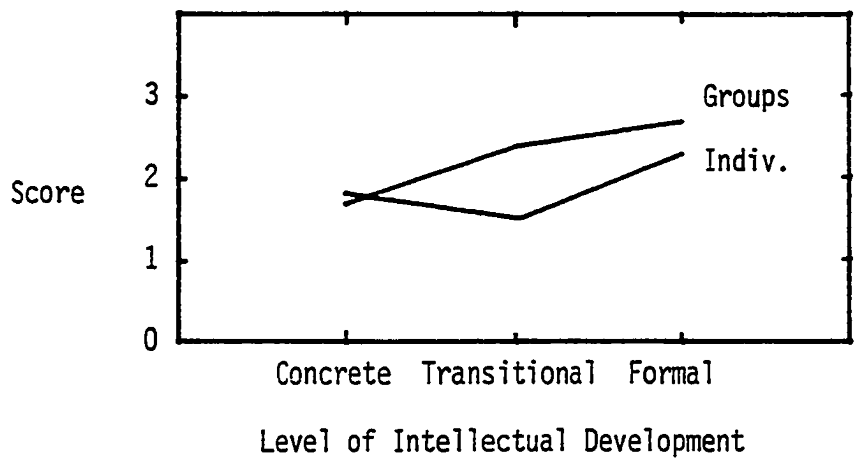
Interaction is another source of variability which is investigated in a two-way ANOVA. This source indicates whether the level of intellectual development on problem solving scores depend on students solving the problems individually or in groups. The cell means data from Tables 10 and 11 are graphed (Graphs I and II) and help indicate interaction effects.

From Table 12, the F-ratio for the interaction effect was significant at $\alpha = 0.20$ for the Aquarium problem (critical $F^{0.20}_{2,60} = 1.65$) but not significant for $\alpha = 0.25$ for the Hills problem (critical $F^{0.25}_{2,60} = 1.42$). From Graph II it can be seen that the differences in the Hills problem scores between groups and individuals is about the same within each level of intellectual development. At least these differences are not significant at $\alpha = 0.25$. Therefore, the levels of intellectual development of the students do not affect the problem scores received by groups or individuals. However, the differences in the scores on the Aquarium problem between groups and individuals does depend on the level of intellectual development. Concrete operational students answering the Aquarium problem individually received slightly higher scores than students at the same intellectual level responding in groups. On the other hand students at the transitional intellectual level answering the problems individually received much lower scores than transitional students responding in groups. Similarly, formal operational individuals received

Graph I

Cell Means - 2-way ANOVA - Group versus Individuals, both Without Previous Group Experience and Intellectual Level

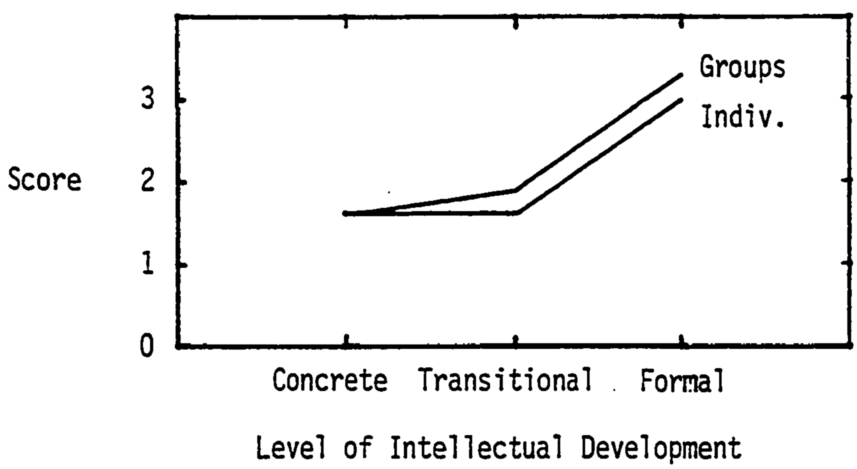
Aquarium Problem



Graph II

Cell Means - 2-way ANOVA - Group versus Individuals, both Without Previous Group Experience and Intellectual Level

Hills Problem



lower scores than formal operational groups, but the difference between groups and individuals was not as great as between transitional groups and individuals. Therefore, the Aquarium problem solving scores received by groups and individuals depended on the levels of intellectual development of the students.

Group and individual performances were compared by disregarding the grade level classification of the students and carrying out a one-tailed t-test. The results of this analysis for each problem are shown in Tables 14 and 15, respectively.

TABLE 14

t-TEST FOR GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE--AQUARIUM PROBLEM

	Mean	Variance	D.F.	t
Groups	2.09	1.09	64	1.60[1]
Individuals	1.70	0.91		

[1] $p<0.1$

TABLE 15

t-TEST FOR GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE--HILLS PROBLEM

	Mean	Variance	D.F.	t
Groups	1.91	0.77	64	0.80[1]
Individuals	1.73	0.96		

[1] $p<0.25$

For both problems the mean performance of the groups was higher than that of the individuals. The obtained t-value for the Aquarium problem was significant at $\alpha = 0.1$ (critical $t_{60}^{0.1} = 1.296$). Significance was not, however, reached for the Hills problem until $\alpha = 0.25$ (critical $t_{60}^{0.25} = 0.679$). These tests have a power of 98 per cent at $\alpha = 0.05$ with a standard deviation difference of $1\ \sigma_e$.

Consider the relevance of the two-way ANOVA and the t-test to the classroom situation. Teachers deal with students at varying levels of intellectual development within the one class. Normally teachers do not provide separate courses of study for students at different levels of intellectual development within the one class. This variable is usually not treated separately in pedagogical practices. If the teacher is considering implementing group work in the classroom the following question is likely to be asked for a particular grade: Will group work lead to higher quality solutions to problems than individual work? It is not until this question is answered that the teacher will consider whether intellectually homogeneous groups provide higher or lower quality responses than intellectually heterogeneous groups.

Because teachers deal with separate grades, the most pertinent analysis from the teacher's point of view would be a two-way ANOVA with the independent variables of groups without previous group experience against individuals without previous group experience and grade, rather than an analysis which controls for level of intellectual development. Hypotheses one and six were tested using this analysis. Data were collected from groups without previous group experience and individuals without previous group experience from grades ten, eleven and twelve. All of the groups and individuals without previous group experience in each grade were used. In addition to testing groups against individuals, this analysis permitted the testing of hypotheses three and eight which were concerned with the affect of grade on problem solving performances.

A two-way ANOVA with unequal cell sizes was used to test the hypotheses. Because this ANOVA was non-orthogonal, the hypotheses were modified so that an analysis which provided independent sums of squares could be used. (See Appendix I for a discussion of non-orthogonal ANOVA's.) If the original hypotheses had been tested, the sums of the squares for the analysis would not have been independent. This could have led to sources of variability being lost or hidden, because of the analytical procedures used to test the hypotheses. It was considered that this was undesirable. Therefore, the following modified null hypotheses were tested.

H_o:1(a). For students without previous group experience, the effect of groups versus individuals will not explain a significant amount of the variability in the scores of the Aquarium problem, ignoring the effects of grade level and interaction.

H_o: 3(a). The effect of grade level will not explain a significant amount of the variability in the scores on the Aquarium problem, adjusting for the effect of groups without previous group experience versus individuals without previous group experience and ignoring interactions.

H_o: 6(a) and H_o: 8(a). Hypotheses similar to H_o:1(a) and H_o: 3(a) were proposed, except that they were concerned with the Hills problem.

These hypotheses were tested using the computer program in Statistical Package for the Social Sciences, Analysis of Variance and Covariance Option 10 (Nie et al., 1975). The cell means for both problems are given in Tables 16 and 17.

Consider the data in Tables 16 and 17. For each grade, except the eleventh graders answering the Hills problem, the mean scores of the groups were higher than the individuals. The data in the tables show that the cell means for the twelfth grade students were not necessarily greater than the means for the tenth and eleventh grade students. In fact, the tenth grade students working in groups scored as high on the Hills problem as the twelfth grade student groups. This trend was not evident in the mean scores for the Aquarium problem, even though the same groups of students responded to both problems. With the Aquarium problem, the twelfth grade student groups performed better than the tenth grade student groups. Only with the individuals solving

TABLE 16

CELL MEANS--2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 10, 11, 12--AQUARIUM PROBLEM

Grade	Groups	Individuals
10	2.2	1.6
11	2.1	1.9
12	2.7	1.7

TABLE 17

CELL MEANS--2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 10, 11, 12--HILLS PROBLEM

Grade	Groups	Individuals
10	2.1	1.5
11	1.7	1.7
12	2.1	1.8

the Hills problem was there evidence of the trend of students at higher grades performing better than students at lower grades.

Summary tables of the ANOVA tests are given in Tables 13 and 14, respectively.

TABLE 13

SUMMARY TABLE: 2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 10, 11, 12--AQUARIUM PROBLEM

Source	SS	df	MS	F
Group vs. Individual	10.30	1	10.30	9.87$^{(1)}$
Grade	2.06	2	1.03	0.99$^{(2)}$
Interaction	2.97	2	1.49	1.42$^{(3)}$
Error	131.59	126	1.04	
Total	146.92	131		

$^{(1)}p<.005$ $^{(2)}p>0.25$ $^{(3)}p<0.25$

TABLE 14

SUMMARY TABLE: 2-WAY ANOVA--GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AGAINST INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 10, 11, 12--HILLS PROBLEM

Source	SS	df	MS	F
Group vs. Individual	1.36	1	1.36	1.68$^{(1)}$
Grade	1.25	2	0.62	0.77$^{(2)}$
Interaction	2.13	2	1.07	1.31$^{(3)}$
Error	101.44	125	0.81	
Total	106.18	130		

$^{(1)}p<0.20$ $^{(2)}p>0.25$ $^{(3)}p>0.25$

The groups against individuals F-ratio was significant at $\alpha = 0.005$ for the Aquarium problem (critical $F^{\alpha=0.005}_{1,120}$ = 8.18), but the F-ratio for the Hills problem was significant at 0.2 (critical $F^{\alpha=0.2}_{1,120} = 1.66$). These ANOVA tests had a power of 97.6 per cent for the group versus individual variable at $\alpha = 0.05$ and a difference of 0.7 σ_e. The rejection of the null hypotheses, H_o: 1(a) and H_o: 6(a) will be discussed in Chapter V. If the null hypotheses are rejected, then a significant amount of the variability in the scores for the Aquarium and Hills problems can be explained by the independent variable of groups versus individuals, both without previous group experience. These results suggest that by forming groups for the first time, students are able to produce higher quality solutions to problems than individuals.

Consider the other variable, grade. The F-ratios for both problems were not significant at $\alpha = 0.25$ (critical $F^{\alpha=0.25}_{2,120} = 1.40$). If neither of the null hypotheses H_o: 3(a) and H_o: 8(a) are rejected, these results suggest that a significant of the variability in the scores for the Aquarium and the Hills problem cannot be explained by the independent variable, grade level.

The interaction F-ratio for the Aquarium problem just reaches significance at $\alpha = 0.25$ (critical $F^{0.25}_{2,120} = 1.40$). Significance for the F-ratio for the Hills problem is not reached at $\alpha = 0.25$. Hence, the problem solving scores of the groups or individuals is affected very little, if at all, by grade.

Hypotheses 4 and 9

The hypotheses four and nine were to be tested using a two-way ANOVA with the independent variables of intellectually homogeneous groups against intellectually heterogeneous groups, both without previous group experience, and grades ten, eleven and twelve. Such a test required cell sizes of nine to achieve a power of 95 per cent with a standard deviation of 1 σ_e. It can be seen from Table 6 that after the data from West Mid-High School were excluded only two heterogeneous groups at grade ten and six homogeneous groups at grade 12 remained. There were, therefore, insufficient data to carry out the analysis as originally planned. Reasons for this insufficiency are discussed in the second section of this chapter. The tenth grade data were dropped from the analysis and a two-way ANOVA was performed with the independent variables of intellectually homogeneous groups against intellectually heterogeneous groups and grade levels eleven and twelve. Neither group had previous group experience. Because all available data were used, unequal cell sizes resulted. This required a non-orthogonal test.

It has already been discussed why a non-orthogonal ANOVA requires a modification in the hypotheses. Hypotheses 4 and 9 were modified as follows:

H_o: 4(a). The independent variable of degree of intellectual heterogeneity will not explain a significant amount of the variability in the scores on the Aquarium problem, ignoring the effects of the variable of grade level and interaction.

H_o:9(a). A hypothesis similar to H_o:4(a) was proposed, except it dealt with the Hills problem.

For a discussion of non-orthogonal ANOVA, see Appendix I.

Cell means for both problems are given in Tables 20 and 21, respectively. Summary tables for the two-way ANOVAs are given in Tables 22 and 23, respectively.

Consider the data in Tables 20 and 21. Intellectually heterogeneous groups had higher mean scores than the intellectually homogeneous groups for the Aquarium problem. This pattern of results was not obtained for the Hills problem. For eleventh grade students, intellectually heterogeneous and homogeneous groups had the same mean score. Twelfth grade homogeneous groups had a higher mean score than heterogeneous groups.

The F-ratio for the homogeneous-heterogeneous groups was significant at $\alpha = 0.25$ for the Aquarium problem (critical $F^{\alpha=.25}_{1,40} = 1.36$) but significance was reached after $\alpha = 0.25$ for the Hills problem. There was a power of 91 per cent at $\alpha = 0.05$ and a difference of 1 σ_e for these tests.

Reasons for not rejecting these null hypotheses

TABLE 20

CELL MEANS--HOMOGENEOUS AGAINST HETEROGENEOUS GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 11 AND 12--AQUARIUM PROBLEM

Grade	Homogeneous	Heterogeneous
11	2.0	2.2
12	2.5	2.8

TABLE 21

CELL MEANS--HOMOGENEOUS AGAINST HETEROGENEOUS GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 11 AND 12--HILLS PROBLEM

Grade	Homogeneous	Heterogeneous
11	1.7	1.7
12	2.3	1.9

TABLE 22

SUMMARY TABLE: 2-WAY ANOVA--HOMOGENEOUS AGAINST HETEROGENEOUS GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 11 AND 12--AQUARIUM PROBLEM

Source	SS	df	MS	F
Homogeneous/Heterogeneous	1.54	1	1.54	1.45 (1)
Grade	2.81	1	2.81	2.63 (2)
Interaction	0.00	1	0.00	0.00 (3)
Error	48.06	45	1.07	
Total	52.41	48		

(1) $p < 0.25$ (2) $p < 0.2$ (3) $p > 0.25$

TABLE 23

SUMMARY TABLE: 2-WAY ANOVA--HOMOGENEOUS AGAINST HETEROGENEOUS GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AND GRADES 11 and 12--HILLS PROBLEM

Source	SS	df	MS	F
Homogeneous/Heterogeneous	0.01	1	0.01	0.01 (1)
Grade	1.80	1	1.80	1.89 (2)
Interaction	0.51	1	0.51	0.54 (3)
Error	43.02	45	0.96	
Total	45.34	48		

(1) $p > 0.25$ (2) $p < 0.2$ (3) $p > 0.25$

are discussed in Chapter V. If the null hypotheses are not rejected, then the results indicate that a significant amount of the variability in the scores received on either problem is not explained by the degree of intellectual heterogeneity within the group. It does not make any difference to the variability in scores received by groups without previous group experience whether these groups are intellectually homogeneous or intellectually heterogeneous.

The results of the independent variable grades, eleven and twelve, were also part of the information received from the two-way ANOVA. For both the Aquarium and Hills problems, the F-ratios were significant at $\alpha = 0.20$ (critical $F^{0.20}_{1,40} = 1.70$). These results suggest that grade level does not explain a significant amount of the variability in the scores received by students. The variability in the scores received by students in grade twelve is not significantly higher than those received by students in grade eleven.

Interaction is another source of variability. Results with this source of variability indicate whether the variability in the scores received by heterogeneous and homogeneous groups depends on the grade level of the students. The interaction F-ratios for both problems were not significant at $\alpha = 0.25$ (critical $F^{0.25}_{1,40} = 1.36$). These results suggest that the differences in the variability of the scores received by homogeneous and heterogeneous groups, without previous group experience, are not related to the different grades of the student.

Hypotheses 5 and 10

Hypotheses five and ten were investigated. These hypotheses are:

H_1: 5. The mean score on the Aquarium problem obtained by tenth grade students who are experiencing a biology class based on group experience will be higher than the mean score obtained by tenth grade students who are experiencing biology which is not based on group work.

H_1: 10. This hypothesis is similar to H_1: 5, but is concerned with the Hills problem.

To test hypotheses five and ten a power of 95 per cent at $\alpha = 0.05$ and a standard deviation of 1.5 σ_e would have required a cell size of thirteen. The test compared tenth grade groups with previous group experience against tenth grade groups without previous group experience. Again, the data in Table 6 show that the analysis with cell sizes of thirteen could not be carried out as planned. Reasons for the insufficiency are given in the second section of Chapter IV. Only eight groups were available at Moore Central Mid-High School (school No. 5). A one-tailed t-test with all the groups from West Mid-High School and Moore Central Mid-High School was used to test the hypotheses. The Aspin-Welch correction to the degrees of freedom was used, because of the unequal cell sizes. This correction is given by the following equation (Myers, 1972, p. 73):

$$df = \frac{(n_1-1)(n_2-1)}{(n_2-1)C^2 + (n_1-1)(1-C)^2}$$

where: $$C = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2}$$

s_j^2 and n_j are the variance and sample size for the j^{th} treatment group.

The results of this analysis for the Aquarium and Hills problems are shown in Tables 24 and 25.

TABLE 24

t-TEST OF PREVIOUS GROUP EXPERIENCE AGAINST NO PREVIOUS GROUP EXPERIENCE FOR TENTH GRADE STUDENTS--AQUARIUM PROBLEM

	Mean	Variance	Corrected d.f.	t
Group Experience	3.07	0.84	10	2.08(1)
No Group Experience	2.00	2.29		

(1) $p < .05$

TABLE 25

t-TEST OF PREVIOUS GROUP EXPERIENCE AGAINST NO PREVIOUS GROUP EXPERIENCE FOR TENTH GRADE STUDENTS--HILLS PROBLEM

	Mean	Variance	Corrected d.f.	t
Group Experience	2.71	1.45	20	1.28(1)
No Group Experience	2.13	0.41		

(1) $p < 0.15$

The data in Tables 24 and 25 indicate that, for both problems, the means of the groups with previous group experience was greater than the mean of the groups without previous group experience. Furthermore, the obtained t-value for the Aquarium problem was greater than the critical t-value at $\alpha = 0.05$ (critical $t_{10}^{.05} = 1.812$). However, the obtained t-value for the Hills problem was greater than the critical value at $\alpha = 0.15$ (critical $t_{20}^{0.15} = 1.064$). The power of this statistical test was 93 per cent for a difference of 1.5 σ_e at $\alpha = 0.05$.

Reasons for rejecting the null hypotheses for the two problems are given in Chapter V. Rejection of the null hypotheses would imply that previous group experience does affect the scores received by the groups. Previous group experience may help the groups function more effectively and thereby produce higher quality solutions than groups without prior group experiences.

General Interpretation and Additional Hypothesis Testing

The major purpose of the present study was to investigate the differences in the quality of responses given by groups and individuals to science problems. Three different analyses were used to test the differences between groups and individuals solving problems.

As discussed earlier in this chapter, the decisions to reject null hypotheses will be made in Chapter V after consideration of the educational implications of such decisions. Therefore, the interpretations of the data given in this section will be based only on descriptive statistics, rather than inferential statistics.

The analyses used to investigate groups and individuals solving science problems were two-way ANOVAs and a t-test. The cell means for the ANOVAs are given in Tables 10, 11, 16 and 17. Tables 14 and 15 give the means for the t-test. The differences between the mean scores of the groups without group experience and individuals without group experience were obtained from each of these tables. These data are given in Table 26.

TABLE 26

DIFFERENCES IN MEAN SCORES BETWEEN GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE AND INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE FROM THREE ANALYSES--AQUARIUM AND HILLS PROBLEMS

Group/Individual Analysis	Problem	
	Aquarium	Hills
and intellectual level (Tables 10 and 11)		
Concrete	-0.1	0.0
Transitional	0.9	0.3
Formal	0.4	0.3
and grade (Tables 16 and 17)		
10	0.6	0.6
11	0.2	0.0
12	1.0	0.3
t-test (Tables 14 and 15)	0.39	0.18

The performances of groups without previous group experience were higher than those of individuals without previous group experiences in all cases except two, Eleventh grade individuals answered the Hills problem better than eleventh grade groups and individuals at the concrete intellectual level responded better to the Aquarium problem than intellectually homogeneous-concrete groups. The largest difference of 0.93 occurred with the intellectually homogeneous-transitional groups. These results suggest that just forming a group, even though members of that group have not worked together before, may lead to the group producing higher quality solutions than students solving the problems alone.

Now consider the differences between the mean scores of groups with previous group experience and groups without previous group experience. Data in Tables 24 and 25 were used to compute these mean differences which are presented in Table 27.

TABLE 27

DIFFERENCES IN MEAN SCORES BETWEEN GROUPS WITH PREVIOUS GROUP EXPERIENCE AND GROUPS WITHOUT PREVIOUS GROUP EXPERIENCE--AQUARIUM AND HILLS PROBLEMS

Aquarium Problem	Hills Problem
1.1	0.6

The Aquarium problem difference of 1.1 between the mean scores of the groups with previous group experience and groups without previous group experience represented a 28 per cent change, because the scores ranged from zero to four. Furthermore, this was the largest difference found amongst all the cell means in Tables 26 and 27. These data suggest that differences in the ability to solve the problem may occur between groups with previous group experience and individuals without previous group experience. This interpretation led to the following alternative hypotheses being postulated and investigated:

H_1: 11. The mean score on the Aquarium problem obtained by groups of tenth grade students with previous group experience will be higher than that obtained by tenth grade students without previous group experience who solve the problem individually.

H_1: 12. A hypothesis similar to the one above was proposed, but it was concerned with the Hills problem rather than the Aquarium problem.

Again, all of the relevant group data available was used for the study. There were fourteen groups from the tenth grade at West Mid-High School. Data for the individuals without previous group experience came from all of those tenth grade students who were part of the first two-way ANOVA performed to test hypotheses two, six and seven. A one-tailed t-test was used to test the hypotheses eleven and twelve. Because the cell sizes were unequal, the Aspin-Welch correction was applied to the degrees of freedom. The results of the tests for the Aquarium and Hills problems are shown in Tables 28 and 29.

TABLE 28

t-TEST OF GROUPS WITH PREVIOUS GROUP EXPERIENCE VERSUS INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE--TENTH GRADE STUDENTS--AQUARIUM PROBLEM

	Mean	Variance	Corrected d.f.	t
Groups with Previous Group Experience	3.07	0.84	6	3.93[1]
Individuals without Previous Group Experience	1.56	0.78		

[1] $p<0.005$

TABLE 29

t-TEST OF GROUPS WITH PREVIOUS GROUP EXPERIENCE VERSUS INDIVIDUALS WITHOUT PREVIOUS GROUP EXPERIENCE--TENTH GRADE STUDENTS--HILLS PROBLEM

	Mean	Variance	Corrected d.f.	t
Groups with Previous Group Experience	2.71	1.45	20	2.72[1]
Individuals without Previous Group Experience	1.44	0.78		

[1] $p<0.01$

The obtained t-value for the Aquarium problem is significant at $\alpha = 0.005$ (critical $t_6^{0.005} = 3.707$), whilst for the Hills problem the obtained t-value is significant at $\alpha = 0.01$ (critical $t_{20}^{0.01} = 2.528$). The power of this test was 90 per cent at $\alpha = 0.05$ for a difference of $1.0\ \sigma_e$.

If the null hypotheses for the two problems are rejected, these results suggest that the previous group experience of the groups enabled the groups to receive higher scores than the students solving the problems alone. Reasons for rejecting the null hypotheses are discussed in Chapter V.

The higher mean scores of groups with previous group experience over individuals without previous group experience might be explained by the students with previous group experience having higher levels of intellectual development than the individuals without group experience. This possible explanation arose because the data presented in Tables 10 and 11 suggested that students at higher levels of intellectual development tend to receive higher scores than students at lower levels. Therefore, the following two additional alternative hypotheses (H_1: 13 and H_1: 14) were proposed and tested.

H_1: 13. There will be a difference in the level of intellectual development between the tenth grade students with previous group experience who answered the Aquarium problem in groups and those tenth grade students without previous group experience who answered the problem individually.

H_1: 14. A hypothesis similar to H_1: 13 was proposed except that H_1: 14 was concerned with the Hills problem.

Ideally, to test these hypotheses, a two-way ANOVA with the main effects of groups with previous group experience against individuals without previous group experience and levels of intellectual development should be used. Such an analysis, however, would require using only the homogeneous-concrete, transitional and formal groups from West Mid-High School. From Table 6 it can be seen that there was no homogeneous-formal group, only three homogeneous-transitional and two homogeneous-concrete groups. This was insufficient data for the analysis. Instead, a two-tailed t-test for each problem was used to compare the mean intellectual level of students with previous group experience, who had solved the problems in groups with those tenth grade students, without previous group experience, who had solved the problems individually.

There were nine students, without previous group experience, from the tenth grade who had solved the problems individually and who were part of the data used to test hypotheses one and six. The level of intellectual development of each of these students was available from the Piagetian tasks. There were, however, fourteen groups each of four students which produced the groups--with previous group experience data. A total of fifty-six students were involved. Rather than using all fifty-six students, determining their levels of intellectual development and comparing them with the nine individuals without previous group experience, a random sample of nine students was selected. This procedure allowed for a test with equal cell sizes to be performed. If the intellectual levels of all fifty-six students had been compared with the intellectual levels of only nine students, a very powerful t-test would have resulted. By using cell sizes of nine, the power of the t-test was 85 per cent at $\alpha = 0.05$ for a standard deviation difference of 1.5 σ_e.

The results of the t-tests for each problem are shown in Tables 30 and 31. The obtained t values for the Aquarium or the Hills problems were not significant with $\alpha = 0.6$ (critical $t_{16}^{0.6} = 0.535$). Rejection of the two null hypotheses would suggest that there are no significant differences between the mean scores of the students who solved the problems in groups and those students who solved the problems alone. The decisions to reject the null hypotheses are discussed in Chapter V.

Hence, the differences that were obtained in testing hypotheses eleven and twelve cannot be explained by differences in the intellectual levels of the students in the groups with previous group experience and the individuals without previous group experience. It is hypothesized that the differences are due to the group experience gained by the West Mid-High School students whilst studying a small group-oriented, laboratory-oriented biology course.

TABLE 30

t-TEST FOR COMPARING LEVELS OF INTELLECTUAL DEVELOPMENT OF TENTH GRADE STUDENTS IN GROUPS WITH PREVIOUS GROUP EXPERIENCE AND INDIVIDUAL TENTH GRADE STUDENTS WITHOUT PREVIOUS GROUP EXPERIENCE--AQUARIUM PROBLEM

	Mean	Variance	d.f.	t
Students in Groups with Previous Group Experience	8.53	2.84	16	0.63 (1)
Individual Students without Previous Group Experience	8.00	3.50		

(1) $p<0.6$

TABLE 31

t-TEST FOR COMPARING LEVELS OF INTELLECTUAL DEVELOPMENT OF TENTH GRADE STUDENTS IN GROUPS WITH PREVIOUS GROUP EXPERIENCE AND INDIVIDUAL TENTH GRADE STUDENTS WITHOUT PREVIOUS GROUP EXPERIENCE--HILLS PROBLEM

	Mean	Variance	d.f.	t
Students in Groups with Previous Group Experience	9.79	3.30	16	0.58 (1)
Individual Students without Previous Group Experience	9.33	2.25		

(1) $p<0.6$

Summary of Results

The results of testing the hypotheses in this research are given in Table 32. Levels of significance of the obtained statistic, the critical values and the power of each of the tests are given in this table. The information in Table 32 will be used to help make the decisions to reject or not to reject the null hypotheses. These decisions, discussed in Chapter V, will be based on the educational implications of erroneously rejecting, or not rejecting, the null hypotheses and the probabilities of Type I and Type II errors occurring.

TABLE 32

SUMMARY TABLE OF RESULTS OF HYPOTHESIS TESTING

Hypotheses and Tests	Independent Variables	Power			Aquarium			Hills		
		%	α Level	σ_e	Obt Stat.	Level of Sig	Crit Value	Obt Stat.	Level of Sig	Crit. Value
H:1,H:6 2-Way ANOVA	Group/Indiv. (No prior grp. experience)	98	.05	1	2.71	0.2	1.68	0.77	>0.25	1.35
	Intel. Level	94	.05	1	1.76	0.2	1.65	8.86	0.001	7.76
H:1,H:6 one-tailed t-test	Group.Indiv. (No prior grp. experience)	98	.05	1	1.60	0.1	1.296	0.80	0.25	0.679
H:1(a),H:6(a) 2-way ANOVA	Group.Indiv. (No prior grp. experience)	97.6	.05	0.7	9.87	0.005	8.18	1.68	0.2	1.66
H:3(a),H:8(a)	Grade(10,11,12)	95	.05	0.8	0.99	>0.25	1.40	0.77	>0.25	1.40
H:4(a),H:9(a) 2-way ANOVA	Homogeneous/ Heterogeneous (No prior grp. experience)	91	.05	1	1.45	0.25	1.36	0.01	>0.25	1.36
H:5,H:10 one-tailed t-test	Groups-group exp-groups- no prior grp experience	93	.05	1.5	2.08	0.05	1.812	1.28	0.15	1.064
H:11,H:12 one-tailed t-test	Groups-group exp/Indiv.- no prior grp experience	90	.05	1	3.93	0.005	3.707	2.72	0.01	2.528
H:13,H:14 two-tailed t-test	Intellectual Level (10th gr. students)	85	.05	1.5	0.63	0.60	0.535	0.58	0.60	0.535

References

Kirk, R. E. Experimental Design: Procedures for the Behavioral Sciences. Belmont, California: Brooks/Cole Pub. Co., 1968.

Myers, J. L. Fundamentals of Experimental Design. 2nd ed. Boston: Allyn and Bacon, Inc., 1972.

Nie, N. H., et al. Statistical Packages for the Social Sciences. New York: McGraw-Hill Book Company, 1975.

Pearson, E. S. and Hartley, H. O. Charts of the Power Function for Analysis of Variance Tests Derived from the Non-Central F-Distribution. Biometrika, 1951, 38.

Renner, J. W. Evaluating Intellectual Development Using Written Responses to Selected Science Problems. The Story of the Cognitive Analysis Project. The University of Oklahoma, 1977.

Renner, J. W. et al. Research, Teaching and Learning with the Piaget Model. Norman: The University of Oklahoma Press, 1976.

CHAPTER V

CONCLUSIONS, EDUCATIONAL IMPLICATIONS AND RECOMMENDATIONS OF THE RESEARCH

Conclusions and Educational Implications

The conclusions of the present research will be based on the decisions to reject or not to reject the null hypotheses. These decisions will be determined by the α-value at which the particular statistic for a given statistical test is significant and the educational implications of falsely rejecting or falsely not rejecting the null hypothesis at that level of significance.

Hypotheses 1, 1(a), 6 and 6(a) (Groups against Individuals--No Previous Group Experience)

A comparison of the quality of the solutions to two science problems given by groups and individuals, both without previous group experience, was of primary concern in this study and was investigated through hypotheses one, one (a), six and six (a). Three different analyses were used to test these hypotheses. Irrespective of the particular analysis, the implications of rejecting or not rejecting the null hypotheses will be the same.

If the decision is made to reject the null hypothesis, then the conclusion of this research would be that there is a difference between groups and individuals, both without group experience, when they solve the problems. Teachers implementing this conclusion would organize students to solve problems in groups rather than individually. Suppose that the conclusion of the research is incorrect. In this instance, students in schools would become involved in group work from which the benefit could be minimal. Students may not, however, be harmed educationally. Hence, a Type I error, falsely rejecting the null hypothesis, can be tolerated in this research.

On the other hand, if the decision is made not to reject the null hypothesis, teachers would not involve students in group work. In this case, an incorrect decision not to reject the null hypotheses would mean that students would not have the opportunity to benefit from solving problems in groups. This would be to the students' educational detriment. Therefore, the error of falsely not rejecting the null hypothesis, a Type II error, should be minimized.

Consider the data in rows one, three and four in Table 32. These data indicate that the four F-ratios and the two t-statistics for the Aquarium and Hills problems are significant at α-values ranging from 0.005 to more than 0.25. The probabilities of falsely rejecting the null hypotheses also range from 0.005 to more than 0.25. The decision was made to accept probabilities of 0.25 or less. The powers of these tests (Table 32) are 98 per cent and 97.6 per cent to detect differences of 1 σ_e and 0.7 σ_e, respectively, at $\alpha = 0.05$. The null hypotheses one, one (a), six and six (a) will be rejected in all tests except one. The probabilities that these hypotheses have been incorrectly rejected range from 0.005 to 0.25 per cent. Hypothesis six, however, which is associated with the Hills problem and the first two-way ANOVA with the independent variables groups against individuals, both without previous group experience, and intellectual level has an obtained F-ratio, $F_{1,60}$ of 0.77. The level of significance of this ratio cannot be obtained from statistical tables and hence the probability of falsely rejecting the null hypotheses cannot be determined. Therefore, the null hypothesis from the first two-way ANOVA that there is no difference in the mean scores on the Hills problem of groups and individuals, both without previous group experience, will not be rejected. If the null hypotheses are not rejected at α-values greater than 5 per cent, the probabilities of Type II errors are less than 2 per cent or less than 2.4 per cent. These probabilities would be within acceptable limits.

Hypotheses 2 and 7 (Level of Intellectual Development)

Contrary to hypotheses one and six, the decision to reject or not to reject hypotheses two and seven will be determined by the probability of a Type I error.

If the null hypotheses are rejected, the conclusion would be made that the mean scores received on the problems do depend on the students' levels of intellectual development. Because the problems were developed with the aim of requiring the use of the higher mental processes characteristic of intellectual maturity, this conclusion would provide evidence for the statement that the aim was fulfilled. The problems and their scoring hierarchies might then be useful instruments for additional research into the relationship between students' levels of intellectual development and the quality of responses to the problems.

The null hypotheses may, however, be rejected when they are true. To erroneously reject a null hypothesis when it is true, leads to a Type I error. If this error were made, contrary to what was believed to be correct, problem scores would not depend on intellectual level. As stated earlier, further research may use these problems and the scoring hierarchies.

The present research could be based on the incorrect premise that the problem scores are dependent upon intellectual level. It would be desirable to avoid this situation. Therefore, Type I errors should be minimized.

The null hypotheses may, however, not be rejected. The conclusion would follow that the mean scores on the problems do not depend on intellectual development. Therefore, the aim would not be realized that the problems and scoring hierarchies should depend on the students' levels of intellectual development. If further research were to be carried out which investigated the relationship between the quality of students' responses and intellectual level, then the Aquarium and Hills problems would not be recommended to be used. Other problems would have to be devised. Alternatively future research might investigate what mental abilities, if any, may be represented in the different qualities of responses given by the students to the Aquarium and Hills problems.

A Type II error may, however, be made by incorrectly not rejecting the null hypothesis when in fact it is false. Consequently the problem scores would really be related to level of intellectual development even though the conclusion from the research would not state this relationship. Therefore, the problems could be used in further research and it would be unnecessary to devise new problems as described in the paragraph above. If the decision not to reject the null hypotheses was indeed incorrect, a useful research instrument, the problems and scoring hierarchies, would be lost. This type of error, Type II, may inhibit future research, but it would not invalidate the research conclusions. Therefore, Type II errors can be tolerated.

In this study, because the test has a power of 94 per cent at $\alpha = 0.05$ and a 1 σ_e difference, the probability of a Type II error at $\alpha = 0.05$ is 6 per cent. The Aquarium problem is significant at $\alpha = 0.2$ and the Hills problem at

$\alpha = 0.1$. Considering the risk of a Type I error, the null hypothesis will be rejected for both problems. Therefore, the mean scores received by students on the problems do not reflect intellectual maturity.

Hypotheses 3(a) and 8(a) (Grades 10, 11, 12)

The hypotheses three (a) and eight (a) were tested with non - orthogonal ANOVA which were not stated in the usual parametric terms. As discussed in Appendix I, the non-orthogonal analyses used in this research required investigation of hypotheses such as whether or not a particular independent variable will explain a significant amount of variability, ignoring the other independent variable and the interaction. Therefore, hypotheses three (a) and eight (a) were not interpreted as whether or not significant differences existed between population means. A conclusion that a significant amount of the variability in the scores received by students for the problem could be explained by grade level would follow from a rejection of hypotheses three (a) and eight (a).

A conclusion that the mean scores depend on the grade level of the students would follow from a rejection of hypotheses three (a) and eight (a). If this conclusion is incorrect, a Type I error would ensure. Teachers implementing the erroneous conclusions of this research would expect students at different grades to give different quality responses, even though students would not be capable of this. These higher expectations of the teachers would be ill-founded. Student capabilities would be judged on incorrect criteria. Such a Type I error, produced by falsely rejecting the null hypothesis, should be minimized. Not to reject the null hypothesis, on the other hand, would lead to the conclusion that problem solutions are not dependent upon grade. A Type II error would result if this decision were incorrect. If teachers implemented an erroneous conclusion in the classroom, they would not expect better quality solutions from students in the higher grades, even though these students would be capable of higher performances than students in the lower grades. The teacher may then consider the possibility of other factors which may be related to solving problems. Students in these higher grades would not be educationally harmed. It could be argued that students live up to the teacher's expectations and that the greater capability of solving problems by students in the higher grades would not be forthcoming. This researcher, however, believes that in many courses taught at the senior levels of high school, teachers have unrealistically high expectations of the intellectual capabilities of the students. Therefore, it is proposed for this research that the students' education will not be hindered by falsely not rejecting the null hypotheses. In other words, a Type II error is not as serious as a Type I error.

The power of the test is 95 per cent at $\alpha = 0.05$

and a 0.8 σ_e difference. The null hypotheses for both the Aquarium and the Hills problems are significant at α-values greater than 0.25. If these hypotheses were rejected, the probability of a Type I error would be greater than 0.25. Such a large error is unacceptable and therefore the null hypotheses will not be rejected. The probability of this decision being incorrect is 5 per cent.

Hypotheses 4(a) and 9(a) (Intellectually Homogeneous and Heterogeneous Groups--No Previous Group Experience)

Hypotheses four (a) and nine (a) were not stated in terms of differences between population means because the hypotheses were tested using non-orthogonal ANOVAs. A rejection of these null hypotheses would lead to the conclusion that a significant amount of the variability in the scores received by students for the problems could be explained by the degree of intellectual heterogeneity within the group.

If the decision to reject the null hypotheses is incorrect, such a conclusion could lead to special intellectual groups being formed when really there is no need. The students' education, however, would not be hindered by working in groups. Therefore, a Type I error is acceptable.

Not to reject the null hypotheses when they should be rejected leads to a Type II error. Special intellectual groups would not be formed, even though students would benefit from them. It is, therefore, desirable to minimize a Type II error.

These tests had a power of 91 per cent at $\alpha = 0.05$ and a difference of 1 σ_e. The probability of rejecting the null hypothesis for the Aquarium problem (H_o: 4(a)), when it is true, is 25 per cent. In other words, the probability of a Type I error occurring is 25 per cent. On the other hand, the probability of a Type II error is less than 9 per cent. This error would result from not rejecting the null hypothesis when it was false. The decision is made not to reject the null hypothesis for the Aquarium problem.

The decision regarding the Hills problem involves additional consequences. The obtained F-ratio for the Hills problem is so low that the level at which this ratio is significant cannot be determined from statistical tables. Therefore, the probability of erroneously rejecting the null hypothesis is not known. The probability, however, of a Type II error is less than 9 per cent. The null hypothesis (H_o: 9(a)) for the Hills problem will not be rejected. Hence, for both the Aquarium and Hills problems, the conclusion is that there are no significant differences between the mean scores of intellectually homogeneous and heterogeneous groups, without previous group experience.

Hypotheses 5 and 10 (10^{th} Grade Groups with Previous Group Experience and Groups without Previous Group Experience)

The decisions to reject the null hypotheses five and ten were based on the need to minimize a Type II error.

The reasons for these decisions are as follows.

An incorrect conclusion that groups with previous group experience perform better than groups without previous group experience would follow from falsely rejecting the null hypotheses. If this erroneous conclusion were implemented in the classroom, students would receive prolonged group instruction from which they would not benefit educationally. The students, however, would not suffer from the experience. Therefore, such a mistake, a Type I error, could be tolerated.

To incorrectly conclude that solving problems does not depend on previous group experience would lead to students not receiving the group instruction from which they could benefit. The education of these students would be hindered. Therefore, it is desirable to minimize the probability of a Type II error.

The probabilities of falsely rejecting the null hypotheses for the Aquarium and Hills problems are 0.05 and 0.15, respectively. Because the power of the tests was 93 per cent at $\alpha = 0.05$ and a 1.5 σ_e difference the probability of falsely not rejecting the null hypotheses is 7 per cent for the Aquarium problem and less than 7 per cent for the Hills problem. The decision was made to reject both null hypotheses and conclude that the mean scores of the tenth grade biology groups with previous group experience were higher than the tenth grade biology groups without previous group experience. Therefore, prolonged group experience does lead to problem responses of a higher quality.

Hypotheses 11 and 12 (10^{th} Grade Groups with Previous Group Experience against 10^{th} Grade Individuals without Previous Group Experience)

Both the null hypotheses eleven and twelve were rejected on the basis of wanting to minimize a Type II error, falsely not rejecting the null hypotheses. The reasons for these decisions are discussed below.

The conclusion that the group performance is better than the individual performance will result from rejection of the null hypothesis. This rejection may be incorrect, a Type I error occur and there reqlly is no difference between groups and individuals. Teachers implementing this erroneous conclusion would provide students with extensive group experience from which they could not benefit. Otherwise, their students' education would not be hindered. Hence, a Type I error can be tolerated.

A decision not to reject the null hypotheses could be wrong and although students could benefit from extensive group experience, teachers will not provide this opportunity. In such a situation, the lack of group experience would be detrimental to the students. Type II errors should, therefore, be minimized.

A power of 90 per cent at an α-value of 0.05 and a standard deviation difference of 1 σ_e was achieved with the tests for hypotheses eleven and twelve. If the null hypotheses are rejected the probability of a Type I error is 0.5 per cent and 1.0 per cent for the Aquarium and Hills problems, respectively. Not to reject the null hypothesis would result in a Type II error of 10 per cent or less for the Hills problem, but more than 10 per cent for the other problem. The decision was made to reject the null hypotheses. This leads to the conclusion that a significant amount of the variability in the scores can be explained by the variable: groups with previous group experience versus individuals without previous group experience. It would appear that prolonged group experience leads to higher quality solutions than those provided by individual students without prior group experience.

Hypotheses 13 and 14 (Intellectual Development--Students in Groups with Previous Group Experience and Individuals without Previous Group Experience)

Hypotheses thirteen and fourteen were tested at a power of 85 per cent, an α-value of 0.05 and a difference of 1.5 σ_e. The data in Table 32 show that both null hypotheses are significant at $\alpha = 0.60$. To reject these null hypotheses could lead to a Type I error. The probability of this error occurring is 0.60. In other words, there is a 60 per cent chance that the null hypotheses were falsely rejected. Because of the high probability of a Type I error, the decision was made not to reject the null hypotheses. This decision could lead to a Type II error--mistakenly not rejecting the null hypotheses when they are false. The probability of a Type II error would be less than 15 per cent. This probability was based on calculations using an α-value of 5 per cent. Not to reject the null hypotheses at α-values greater than 5 per cent will lower the probability of a Type II error to less than 15 per cent.

The decision not to reject the null hypotheses leads to the conclusion that there is no difference in the mean scores of the levels of intellectual development of students, with previous group experience, who solved the problems in groups and students, without previous group experience, who solved the problems individually. There is less than a 15 per cent chance that this decision is incorrect.

The decision not to reject the null hypotheses can also be based on the educational implications of Type I and Type II errors occurring. The other null hypotheses, one through twelve, were rejected or not rejected on the probabilities of Type I and Type II errors occurring and the subsequent educational implications which follow from each of these two types of errors being made. This procedure of considering the implications and probabilities of drawing erroneous conclusions will now be discussed.

First, it should be noted that the decisions whether or not to reject null hypotheses thirteen and fourteen are based on the implications of these decisions on rejecting the null hypotheses eleven and twelve.

Not to reject the null hypotheses thirteen and fourteen would lead to the conclusion that the mean of the intellectual levels of the students with previous group experience is the same as the mean of the intellectual level of students without previous group experience. Such a statement would support the conclusion from hypotheses eleven and twelve that the higher mean scores of the groups with previous group experience was the result of this previous experience, rather than the higher intellectual levels of these students over students without previous group experience. Teachers in schools would then provide group experience for the benefit of students. If there really is a difference in the intellectual levels of the students a Type II error will have occurred. The null hypotheses thirteen and fourteen will not have been rejected when they are false.

A Type II error has implications concerning the conclusions drawn from hypotheses eleven and twelve. These conclusions were that the mean scores of the groups with previous group experience were higher than the individuals, because students worked in groups and because these students had experience planned group work. If Type II errors occurred in the conclusions drawn from hypotheses thirteen and fourteen, the conclusions following hypotheses eleven and twelve would be incorrect. Previous group experience would, in fact, not necessarily have resulted in higher quality solutions of the groups with prior group experience compared with the individuals without previous group experience. The differences in the scores between groups and individuals could have been caused by the students in the groups having higher levels of intellectual development than the students who solved the problems individually. Therefore, teachers who had implemented the conclusions from hypotheses eleven and twelve would be providing planned group experiences for students based on the erroneous belief that these experiences would lead to groups of students producing higher quality solutions to problems. The quality of the solutions would in fact, be partly due to the students' levels of intellectual development. Therefore, the anticipated benefit of organized group experiences could be minimal. No additional hindrance to these students' education is likely to result. Consequently, Type II errors can be accepted.

To reject the null hypotheses thirteen and fourteen, and conclude that there is a difference in the intellectual levels of the students, would imply that the observed differences between the mean scores of the groups with previous group experience and the individuals without group experience was not the result of the group experience, but the result of differences in the intellectual levels of the students. This would place in doubt the conclusion that the organized group experiences lead to groups receiving higher scores than individuals. It would be desirable not to disregard this conclusion, because teachers would not provide group work which may benefit the students. Therefore, the error, of rejecting the null hypotheses thirteen and fourteen when they are true, should be avoided, because this has undesirable educational implications on the conclusions drawn from hypotheses eleven and twelve. Hence, Type I errors should be reduced for hypotheses thirteen and fourteen.

General Conclusions and Educational Implications

It would appear that groups of tenth grade biology students, with previous group experience, solved the two science problems better than tenth grade biology students, without previous group experience, who solved the problems individually. This difference in the performances could not be explained by differences in the intellectual levels of the students. Furthermore, group experience over an extended period of time also leads to academic performances superior to that of groups which have been formed for the first time. Students working in groups for the first time, who have not had planned group experiences, performed better than individuals, without any group experience. These data suggest that, not only do groups formed for the first time provide higher quality solutions than individuals, but also that planned group experience will enhance the academic performance of the group. Therefore, group experience has a cumulative effect on the ability of students to solve science problems.

An educational implication of these conclusions is that not only can group work assist in the organizational problems faced by the teacher but, more importantly, group work can promote the fulfillment of one of the primary areas of education--improving the quality of student responses to problems that require the ability to think.

Moreover, the data suggest that the teacher does not have to be concerned with forming intellectually heterogeneous or homogeneous groups. No one type of group provided higher quality solutions than the other.

A teacher might argue that, although group work may assist in solving problems, little can be achieved unless all teachers provided organized group experiences for the students. The results of this research refute that argument. Students at West Mid-High School received their planned group experiences in just one class, biology, and these students demonstrated a higher ability to solve the problems than comparable students who had not experienced structured group work. These West Mid-High School students had not, however, increased their level of intellectual development above the other comparable tenth grade students. Perhaps, as suggested by Piagetian theory, factors other than group experience (a type of social interaction) are required to promote intellectual development.

When the two problems were selected for use in this study, it was hoped that these problems would require the use of the higher intellectual abilities characteristic of intellectual maturity. This hope was not realized for either problem. Perhaps the problems only elicited responses at the concrete level of intellectual development. Although formal operational students would be capable of providing a response characteristic of their level of development, the problems did not induce this type of response. Another explanation could be that the scoring hierarchy did not separately categorize the higher quality responses provided by formal operational students.

Grade did not affect the performances of students in their ability to solve the problems. If indeed the quality of the responses to the problems is dependent upon the level of intellectual development of the students, then it is to be expected that grade would not be highly related to the students' performances. Students from only three consecutive grades were used in this study. Differences in grade levels do not necessarily reflect large differences in the intellectual levels of the students. Data presented in Table 5 support this statement. These data demonstrate that the three levels of intellectual development, concrete, transitional and formal are distributed, throughout grades ten, eleven and twelve. Admittedly the percentage of concrete operational students in the tenth grade is higher than that in the twelfth grade, but the difference between those two percentages is only 7.42 per cent. Likewise, the difference in the percentage of formal operational students in the tenth and twelfth grade is only 7.92 per cent. The majority of students in each grade are concrete operational.

Recommendations

The present study only demonstrated the cumulative effect of planned group experience on the quality of responses to two science problems for tenth grade biology students. It is recommended that further investigations be carried out to determine whether organized group experiences can benefit either, students studying other subjects, or students at different grades. Students in grades eleven and twelve working in groups for the first time show academic achievement superior to that of students working individually. Perhaps students in these two grades would also benefit from planned group experiences. If further research is carried out it should be noted that the previous group experience of the tenth grade students was closely controlled and supervised by the teacher and that designing and performing laboratory investigations was an essential aspect of this group experience. In addition, each member of the group had a specific role to fulfill so that the group functioned effectively.

Furthermore, this research only compared groups with and without previous group experience to individuals without previous group experience. Additional research should investigate the relationship between groups and individuals, both with prior group experience and the quality of their responses to problems.

Of the two problems used in this research only one, the Hills, demonstrated a relationship between problem solving scores and the level of intellectual development of the student. Further research could identify those characteristics of problems which are related to intellectual development and those problems which will promote intellectual development. This would enable the evaluation of students to be based on their level of intellectual development, rather than their ability to memorize and recall factual information. Promotion of the ability to think, through intellectual development, would thereby fulfill the central purpose of American education.

APPENDIX A

APPENDIX A

THE AQUARIUM PROBLEM

NAMES ______________________ SCHOOL ______________________

____

____

____

BIRTHDATE __________________ GRADE ______________________

When an aquarium is set up the sides of the tank are clean and the water is clear. Within a few months the water and the sides of the aquarium become green. The aquarium has water, fish, plants and gravel in it. The aquarium is in a room with ordinary light. Describe the experiment you need to do in order to find out what it is that causes green material to appear in the water and on the sides of the aquarium. Give reasons for what you did.

APPENDIX B

APPENDIX B

THE HILLS PROBLEM

NAMES ____________________ SCHOOL ____________________

____

____

____

BIRTHDATE ________________ GRADE ____________________

Two hills of different heights are found in central Oklahoma. Those hills are full of ruts. The ruts on the big hill are deeper than those on the small hill. At the foot of each hill is found a big deposit of soil. The soil deposit at the foot of the big hill is larger than that at the foot of the small hill. The hills are very close together and are made of the same types of soil. The thickness of the covering of trees, plants and other vegetation growing on the hills is the same.

You are told by a soil expert that the ruts and their depths on the hills are due to running water. That expert also tells you the amount of soil in the soil deposit at the foot of each hill is also due to running water. The only water that flows down the hill comes from rainfall, which is the same on both hills.

Describe the experiments you need to do in order to test whether or not what the soil expert tells you is true.

APPENDIX C

APPENDIX C

A DESCRIPTION OF CENTRAL INNOVATIVE HIGH SCHOOL (OKLAHOMA CITY)

Central Innovative High School is a school established to meet the special needs of secondary education in Oklahoma City. It aims to provide a clear alternative to the learning environment of the traditional high school. The school is based on the philosophy that students want to learn because they are concerned about the quality of their lives.

Not all students will benefit from attending Central Innovative High School. It is believed that students who appreciate the opportunity to explore, assume responsibility for their own lives and their own educational program will benefit most from attending the school. All students who attend the school apply and are selected to represent a cross-section of Oklahoma City Public Schools according to race, sex, home school and grade level. Enrollment in 1977-78 was 340.

There are three phases to the overall program. First, traditional courses such as English and Mathematics are attended by students on a regular basis. Special projects, centered around a theme in which the student has a particular interest or need, form a second aspect of the program. These special projects might include Humanities and Technology which is provided for college-oriented students. Technical and

Communication offers work in photography, video, printing and television. There is also Industrial Arts and Theater Arts. Finally, satellite courses which involve independent study courses comprise the third aspect of the school program. Open laboratories such as language arts, mathematics and reading are provided for students who need additional help. Career awareness programs are offered which may or may not take place in the community. Volunteers, not staff members, offer such courses as karate, German and modern dance.

The school staff is attempting to make the school more open by providing more flexible learning opportunities for students. In the future more options will be offered to the students and more career awareness opportunities will be provided. There will be greater communication between the school and the home so that an individualized educational program can be provided for each student.

APPENDIX D

APPENDIX D

DESCRIPTION OF THE GROUP EXPERIENCE OF STUDENTS AT WEST MID-HIGH SCHOOL

Students in the tenth grade biology class at West Mid-High School gained their group experience by following the Inquiry Role Approach (Bingham et al., 1974; Marek, 1977). This is a program designed to develop inquiry and social skills, understanding of biology content and attitudinal qualities through laboratory investigations, paper-and-pencil problem solving activities and laboratory explorations. During laboratory explorations each team of four students develops and investigates its own problems.

IRA activities introduce students to inquiry and orient the students into group work. To facilitate this small group work each member of the group has a set of responsibilities or roles. One student, the discussion co-ordinator, has the responsibility of directing the team in answering discussion questions. Another student, acting as a technical advisor, leads the team by delegating the work of the experiment. Data gathering and recording are the responsibilities of the data organizer. Any problems of working together or difficulties with the experiment are the concern of the process adviser. As experience in working in groups is gained by the students, the roles become less structured and mechanical.

Approximately one-third of the class time is devoted to working in four-member groups. Individual assignments require about one-third of the time and the remaining time is spent in class discussions of the team work.

References

Marek, E. M. "The Influence of Inquiry Learning on Intellectual Development, Achievement and IQ." Unpublished doctoral dissertation, The University of Oklahoma, Norman, 1977.

Bingham, R. M.; Koutnik, P. G.; Seymour, L. A.; Padberg, L.F.; and Bingham, K. J. Inquiry Role Approach. Morristown, N.J.: Silver Burdett Publishing Co., 1974.

APPENDIX E

APPENDIX E

HIERARCHY PREPARED FOR SCORING THE AQUARIUM PROBLEM AND EXAMPLES OF STUDENT RESPONSES

The Hierarchy

0. No response or answer completely irrelevant or "don't know."
1. An explanation of the results obtained in the problem was given. The students may have described an experiment but it was unrelated to the problem. This was known as "experimental junk." Also an experiment may have been described but it dealt with only 1 of the factors and was not organized, e.g., "take the fish out."
2. A coherent relevant response was given with experimental procedures described which could possibly solve the problem if extended in some way. This category was known as "organized empiricism." It was necessary for the response to deal with an experiment. Sometimes the responses indicated that objects should be successively taken out of the aquarium. This was regarded as an organized response but one which lacked the combinatorial system necessary to solve the problem. Likewise some responses added items successively to the aquarium. Again, these were organized responses but they lacked a system.
3. The solution involved the correct combination of two

elements (e.g., fish and water or plants and water). There was experimental manipulation of one element at a time, with or without water. Similarly the elements may have been manipulated by omitting one of the variables.

4. The response indicated the correct combination of three or more elements in addition to those described under category 3. This represented an addition to those described under category 3. This represented an organized system. All possible combinations of the variables was not necessary. The particular experimental combinations of variables had to be specified not just referred to in general terms such as "set up different combinations."

Examples of Student Responses

Category 0. "The water and plants haven't change, Because you should keep aquarium clean."

Category 1. "Take out the fish and the green plants and leave the water and gravel in. Recoz Aleaga is cozed by fish waste and deteriation oxygen by plants."

Category 2. "First you take the fish away, then the plants, then the gravel, then light, then the water to see if green junk disappear when these are taken away."

Category 3. "Put each of the above listed items in separate aquariums filled w/ water. Put one of each in the dark and one of each in the ordinary room light. The reason we put them individually is to determine which individual is causing the green substance to form. The reason we put them in the light and the dark is to determine whether or not it would grow in light or dar or in both."

Category 4. "There are 4 variables that could have caused the water and aquarium sides to turn green. These variables are fish, plants, and gravel + light. To determine the cause of the green we would place only one of the variables in the aquarium. We will have water and fish in the aquarium and leave it for a few months. We will also do the same for all the other variables combined with water. We are assuming that water by itself wont turn green but if we are to check the water, we could leave it for a few months by itself. The reason for this set up, (of the different variables and water) is to determine what it is (which variable) that causes the water and tank to turn green. Now if none of these combinations achieve the desired results, we will come to the conclusion that the reaction is dependent on more than just one variable.

We would then try combinations like, Gravel, fish + water, light, gravel + water, etc.

APPENDIX F

APPENDIX F

HIERARCHY PREPARED FOR SCORING THE HILLS PROBLEM AND EXAMPLES OF STUDENT RESPONSES

The Hierarchy

0. No response or "don't know" or a completely irrelevant statement.

1. There was an attempt to deal with the problem but it failed. The problem may have been restated. An explanation of the observations or facts given in the problem may have been attempted.

2. A "minimum logical" response represented an organized attempt to solve the problem. The solution was based on empiricism. In other words an experiment was set up but it was not necessarily a controlled experiment. There was some experimental manipulation or operation. The students had to indicate that they were "acting upon" the information given in the problem.

3. An experiment was set up, with controls in mind, to solve the problem. For example, it might be suggested to take measurements before and after rain or to set up an analogous experiment in the laboratory.

4. The experiment described satisfactorily solves the problem and also states the necessary criteria for validating or disproving the hypothesis outlined in the problem. These criteria or conditions have to be stated, not just implied in the response.

Examples of Student Responses

Category 0. "Something to measure with. And a rainfall gage."

Category 1. "Soil experiments to see if the soil is really the same. Experiments to see what really causes the ruts. Experiments to see if the only water is from rainfall."

Category 2. "Just pour water on one of hills and see if running water causes ruts."

Category 3. "take a sediments from each hill (10 pounds) fir each hill you place the two in seperate boxes and level it out. Then set the boxes at a 25 percent grade so the water flows at the same rate. Then let it flow for 10 minutes each hour for 24 hours. Notice any erosion. (The water should be dripped from the same height)".

Category 4. "Measure the ruts and the soil soil deposits of both hills. Cover one hill with rain proof cover so that no rain will fall on it. Leave other hill uncovered. Each month for a year, measure the depth of the ruts and soil deposits of both hills. Determine the amount of change from the findings of each measurement. If rainfall is the only running water, then the ruts and soil deposit of the covered hill should not change in measurements, and the other hills ruts should become deeper and the deposits larger."

APPENDIX G

APPENDIX G

DATA USED IN TESTING THE HYPOTHESES

Hypotheses 1, 2, 6 and 7

Test--2-way ANOVA

Independent Variables--(i) groups without previous group experience against individuals without prior group experience.

(ii) levels of intellectual development (concrete, transitional, formal)

Dependent Variable--problem solving score on (a) the Aquarium problem, (b) the Hills Problem.

Data for the Aquarium Problem

Level of Intellectual Development	Groups				Individuals			
Concrete	1	1	1	2	0	2	1	2
	2	2	2	1	1	2	3	1
	2	2	3	1	4	4	1	1
	2	1	2	2	1	2	1	2
Transitional	4	4	0	3	3	2	1	2
	3	2	1	3	1	1	2	1
	1	3	1	4	1	2	1	1
	2	3			2	1		
Formal	2	4	2		1	3	3	

Data for the Hills Problem

Level of Intellectual Development	Groups				Individuals			
Concrete	2	2	2	2	1	0	3	2
	1	1	2	2	2	1	3	1
	1	1	2	2	1	1	3	2
	1	2	1	2	2	1	1	1
Transitional	2	3	1	2	1	2	2	1
	1	2	2	1	1	1	2	2
	1	1	3	1	4	1	2	2
	3	4			2	0		
Formal	3	4	3		3	4	2	

<u>Hypotheses 1(a), 3(a), 6(a) and 8(a)</u>

Test--2-way ANOVA

Independent Variables--(i) groups without previous group experience against individuals without prior group experience.

(ii) grade level (10,11,12).

Dependent Variable--problem solving score on (a) the Aquarium problem, (b) the Hills problem.

Data for the Aquarium Problem

Grade Level	Groups				Individuals			
10	1	1	1	2	0	1	2	1
	4	4	0	3	2	2	2	1
	2	4	3		1	0	2	1
					1	3	2	1
					2	3	1	1
					2	3		
11	2	2	2	1	2	3	1	1
	2	2	3	1	0	1	4	4
	2	1	3	2	1	1	1	1
	1	3	1	3	1	2	2	3
	4	3	1	3	1	2	2	3
	2	2	4	3	3	3	1	4
	1	1	4	1	1	0	2	2
	1	3	1					
12	2	2	4	2	2	1	1	1
	3	2	1	4	3	1	2	2
	4	4	3	2	1	2	1	1
	4	1	2	2	2	1	1	2
	3	3			2	3	2	3
					1	3		

Data for the Hills Problem

Grade Level	Groups				Individuals			
10	2	2	2	2	1	2	1	1
	2	3	1	2	3	0	1	2
	3	2	2		1	2	1	2
					2	1	1	1
					1	2	2	2
					2			
11	1	1	2	2	2	1	3	0
	1	1	2	2	1	1	1	1
	1	2	1	2	2	1	3	2
	2	1	1	1	1	2	1	2
	3	4	3	3	1	1	4	2
	2	1	0	3	2	1	1	4
	2	1	1	0	2	3	1	2
	2	2	2					
12	1	2	1	3	2	2	1	1
	4	3	2	2	1	1	1	2
	2	3	0	0	2	0	2	2
	2	2	3	2	0	4	3	2
	3	2			2	2	3	2
					3	1		

Hypotheses 4(a) and 9(a)

Test--2-way ANOVA

Independent Variables--(i) intellectually homogeneous groups against intellectually heterogeneous groups.

(ii) grade (11, 12).

Dependent Variable--problem solving score on (a) the Aquarium problem, (b) the Hills problem.

Data for the Aquarium Problem

Grade	Homogeneous				Heterogeneous			
11	2	2	2	1	3	1	3	2
	2	2	3	1	2	4	3	1
	2	1	3	2	1	4	1	1
	1	3	1	3	3			
	1	4						
12	2	2	4	2	1	4	4	4
	3	2			3	2	4	1
					2	2	3	3

Data for the Hills Problem

Grade	Homogeneous				Heterogeneous			
11	1	1	2	2	3	3	2	1
	1	1	2	2	0	3	2	1
	1	2	1	2	1	0	2	2
	2	1	1	1	2			
	3	4						
12	1	2	1	3	2	2	2	3
	4	3			0	0	2	2
					3	2	3	2

Hypotheses 5 and 10

Test--one-tailed t-test

Independent Variable--groups with previous group experience against groups without prior group experience.

Dependent variable--problem solving score on (a) the Aquarium problem and (b) the Hills problem.

Data for the Aquarium Problem

Groups with Group Experience				Groups without Group Experience			
2	4	2	3	1	1	1	4
4	3	4	3	4	0	3	2
3	3	4	4				
3	1						

Data for the Hills Problem

Groups with Group Experience				Groups without Group Experience			
2	4	2	3	2	2	2	2
1	3	0	4	3	1	2	3
2	3	4	3				
3	4						

<u>Hypotheses 11 and 12</u>

Test--one-tailed t-test.

Independent Variable--groups with prior group experience against individuals without prior group experience.

Dependent Variable--problem solving score on (a) the Aquarium problem, (b) the Hills problem.

Data for the Aquarium Problem

Groups with Group Experience				Individuals without Group Experience			
2	4	2	3	0	2	1	2
4	3	4	3	2	1	3	1
3	3	4	4	2			
3	1						

Data for the Hills Problem

Groups with Group Experience				Individuals without Group Experience			
2	4	2	3	2	3	0	1
1	3	0	4	1	2	1	1
2	3	4	3	2			
3	4						

<u>Hypotheses 13 and 14</u>

Test--two-tailed t-test.

Independent Variable--students with previous group experience who answered the problems in groups against students without prior group experience who answered the problems individually. Two separate tests were carried out--one for each problem.

Dependent Variable--levels of intellectual development measured either by the Piagetian task interviews or the CAP measure of intellectual development.

Data for Students who Solved the Aquarium Problem

Students with Group Experience Who Worked in Groups				Students without Group Experience Who Worked Individually			
7.88	6.73	10.76	9.92	5	8	6	8
10.45	9.65	8.01	6.46	8	7	10	11
6.90				9			

Data for Students Who Solved the Hills Problem

Students with Group Experience Who Worked in Groups				Students without Group Experience Who Worked Individually			
7.21	8.13	8.90	13.61	11	8	8	7
10.46	10.38	9.42	9.46	10	9	11	11
10.51				9			

APPENDIX H

APPENDIX H

COMPUTER PROGRAMS

Selection of Random Samples for Individual Data

```
1 PRINT "ro's research  "
2 PRINT "press 1  then 2  then 3  BREAK Twice at end then 4
  good luck"
3 GO TO 16Ø
4 PRINT @37,26:1
5 GO TO 29Ø
8 PRINT "which tape file?"
9 INPUT F
10 FIND F
11 RETURN
12 GO TO 2ØØ
13 GO TO 29Ø
14 FIND F
15 RETURN
16 CLOSE
17 RETURN
20 GO TO 39Ø
21 REM READ DATA OFF TAPE & SORT DATA INTO C/T/F
24 GO TO 116Ø
25 REM READ DATA OFF TAPE--ALL INDIVIDUALS
40 GO TO 69Ø
41 REM  This allows rerandomization w/o rereading tape file.
100 REM Variables are defined as follows:
110 REM  al=student's number
120 REM  gl=PIAGET SCORE
130 REM  standard tape format is al      ,gl
140 REM  in image 4D,2x,3D
150 DIM B$(2Ø),C$(2Ø)
160 B$="STUDENT'S NUMBER?"
170 C$="PIAGET SCORE?"
```

```
180 RETURN
190 IMAGE 4D,2X,3D
200 PAGE
210 PRINT B$
220 INPUT A1
230 PRINT C$
240 INPUT G1
250 GOSUB 32Ø
260 GOSUB 37Ø
270 PRINT " "
280 GO TO 2ØØ
290 PRINT @4Ø:"St. #   PIAGET SC  "
300 PRINT @37,26:Ø
310 RETURN
320 PRINT @37,26:1
330 PRINT @4Ø USING 34Ø:A1,G1
340 IMAGE 4d,15x,3D
350 PRINT @37,26:Ø
360 RETURN
370 PRINT @33: USING 19Ø:A1,G1
380 RETURN
390 PRINT "WHICH FILE?"
400 INPUT F9
410 FIND F9
415 DELETE QØ
420 DIM QØ(2,6ØØ)
430 ON EOF (Ø) THEN 47Ø
440 FOR I=1 TO 6ØØ
450 INPUT @33:QØ(1,I),QØ(2,I)
460 NEXT I
470 I9=I-1
480 DIM C3(3ØØ),T3(3ØØ),F3(3ØØ)
490 J1=1
500 K1=1
510 L1=1
```

```
520 FOR I=1 TO I
530 IF Qø(2,I)<=8 THEN 62ø
540 IF Qø(2,I)<=11 THEN 65ø
550 F3(J1)=Qø(1,I)
560 J1=J1+1
570 NEXT I
580 J1=J1-1
590 K1=K1-1
600 L1=L1-1
610 GO TO 69ø
620 C3(L1)=Qø(1,I)
630 L1=L1+1
640 GO TO 57ø
650 T3(K1)=Qø(1,I)
660 K1=K1+1
670 GO TO 57ø
680 REM RANDOM SAMPLES--ORDER--1ø C:1ø T:1ø F:11 C:11 T:11 F:12 C:
690 DELETE F,N,R
700 DIM F(15ø),N(15ø),R(15ø)
710 F=ø
720 N=ø
730 R=ø
740 Z2=L1-1
750 GOSUB 1ø1ø
760 FOR I=1 TO Z1
770 F(I)=C3(Q1(I))
780 NEXT I
790 Z2=K1-1
800 GOSUB 1ø1ø
810 FOR I=1 TO Z1
820 N(I)=T3(Q1(I))
830 NEXT I
840 Z2=J1-1
850 GOSUB 1ø1ø
860 FOR I=1 TO Z1
```

```
870 R(I)=F3(Q1(I))
880 NEXT I
890 PRINT @37,26:1
900 PRINT @4Ø:"C      T      F "
910 FOR I=1 TO 15Ø
920 IF F(I)+N(I)+R(I)=Ø THEN 96Ø
930 PRINT @4Ø: USING 95Ø:F(I),N(I),R(I)
940 NEXT I
950 IMAGE 3(3D,5X)
960 PRINT @4Ø: USING 97Ø:
970 IMAGE 1ØL
980 PRINT @37,26:Ø
990 END
1000 REM SUBROUTINE:RANDOM SELECTION W/O REPLACEMENT
1010 PRINT "WHAT SAMPLE SIZE?"
1020 INPUT Z1
1030 DELETE Q1
1040 DIM Q1(Z1)
1050 FOR I=1 TO Z1
1060 Q1(I)=INT(Z2*RND(-2))+1
1070 NEXT I
1080 FOR I=1 TO Z1-1
1090 FOR J=I+1 TO Z1
1100 IF Q1(I)=Q1(J) THEN 114Ø
1110 NEXT J
1120 NEXT I
1130 RETURN
1140 Q1(J)=INT(Z2*RND(-2))+1
1150 GO TO 1Ø8Ø
1160 PRINT "WHICH FILE?"
1170 INPUT F9
1180 FIND F9
1190 DELETE QΦ
1200 DIM QΦ(6ØØ)
1210 ON EOF (Φ) THEN 125Ø
```

```
1220 FOR I=1 TO 6ØØ
1230 INPUT @33:QΦ(I),G1
1240 NEXT I
1250 I9=I-1
1260 Z2=I9
1270 GOSUB 1Ø1Ø
1280 DIM W(Z1)
1290 FOR I=1 TO Z1
1300 W(I)=QΦ(Q1(I))
1310 NEXT I
1320 PRINT @37,26:1
1330 PRINT @4Ø:"INDIV"
1340 FOR I=1 TO Z1
1350 PRINT @4Ø: USING 137Ø:W(I)
1360 NEXT I
1370 IMAGE 3D
1380 PRINT @37,26:Ø
1390 END
```

Selection of Random Samples for Group Data

```
1 PRINT "ro's  research  "
2 PRINT "press 1  then 2  then 3  BREAK Twice at end then 4
  good luck"
3 GO TO 19Ø
4 PRINT @37,26:1
5 GO TO 41Ø
8 PRINT "which tape file?"
9 INPUT F
10 FIND F
11 RETURN
12 GO TO 26Ø
13 GO TO 41Ø
14 FIND F
15 RETURN
16 CLOSE
17 RETURN
```

```
20 GO TO 51Ø
21 REM DATA READ OFF TAPE.  Random select homo- c t f
28 GO TO 118Ø
29 REM SORT & RANDOM SELECT--HOMO/HETER;11,12.
32 GO TO 2Ø2Ø
33 REM SORT & RANDOM SELECT-- 1Ø,BIOL;GRP EXP/NO GRP EXP.
36 GO TO 267Ø
37 REM SORT & RANDOM SELECT TOTAL
40 GO TO 294Ø
41 REM Sort and random select  homo-hetero
110 REM Variables are defined as follows:
110 REM SØ=group exper(Ø=no,1=yes)
120 REM gØ=grade, s1=school (2=WMH, 3=MHS, 4=SHS,
130 REM  5=CIHS, 6=CJHS)
140 REM g1=group type (1-3=homogeneous,c,t,f;4-7=heterogeneous,
    c,t,f,m)
150 REM  g2=group number
160 REM  standard tape format is s1,gØ,sØ,g1,g2
170 REM  in image 1d,2x,2d,2x,1d,2x,1d,2x,3d
180 DIM C$(2Ø),J$(2Ø),K$(2Ø),L$(2Ø),M$(2Ø)
190 C$="group experience?"
200 J$="SCHOOL?"
210 K$="GROUP TYPE?"
220 L$="GROUP NUMBER"
230 M$="GRADE?"
240 RETURN
250 IMAGE 1d,2x,2D,2X,1D,2X,3D,2X,3D
260 PAGE
270 PRINT L$
280 INPUT G2
290 PRINT J$
300 INPUT S1
310 PRINT M$
320 INPUT GØ
330 PRINT K$
```

```
340 INPUT G1
350 PRINT C$
360 INPUT S0
370 GOSUB 440
380 GOSUB 490
390 PRINT " "
400 GO TO 260
410 PRINT @40:"group #  school  grade  grp type  grp exp "
420 PRINT @37,26:0
430 RETURN
440 PRINT @37,26:1
450 PRINT @40: USING 460:G2,S1,G0,G1,S0
460 IMAGE 3d,10x,1d,6x,2d,6x,3d,10x,1d,8x
470 PRINT @37,26:0
480 RETURN
490 PRINT @33:G2;S1;G0;G1;S0
500 RETURN
510 PRINT "which file?"
520 INPUT F9
530 FIND F9
540 DELETE Q0
550 DIM Q0(4,125)
560 ON EOF (0) THEN 600
570 FOR I=1 TO 125
580 INPUT @33:Q0(4,I),S1,Q0(3,I),Q0(2,I),Q0(1,I)
590 NEXT I
600 I9=I-1
610 DELETE F
620 DIM C(60),T(60),F(60)
630 J3=1
640 K3=1
650 L3=1
660 FOR I=1 TO I9
670 IF Q0(1,I)=0 THEN 730
680 NEXT I
690 J3=J3-1
700 K3=K3-1
```

```
710 L3=L3-1
720 GO TO 870
730 IF Q0(2,I)=1 THEN 770
740 IF Q0(2,I)=2 THEN 800
750 IF Q0(2,I)=3 THEN 830
760 GO TO 680
770 C(J3)=Q0(4,I)
780 J3=J3+1
790 GO TO 680
800 T(K3)=Q0(4,I)
810 K3=K3+1
820 GO TO 680
830 F(L3)=Q0(4,I)
840 L3=L3+1
850 GO TO 680
860 REM random SAMPLES     -order--    c:    t:    f
870 PRINT "SAMPLE SIZE?"
880 INPUT Z1
890 DELETE F6,G6,H6
900 DIM F6(Z1),G6(Z1),H6(Z1)
910 Z2=J3
920 GOSUB 2530
930 FOR I=1 TO Z1
940 F6(I)=C(Q1(I))
950 NEXT I
960 Z2=K3
970 GOSUB 2530
980 FOR I=1 TO Z1
990 G6(I)=T(Q1(I))
1000 NEXT I
1010 Z2=L3
1020 GOSUB 2530
1030 FOR I=1 TO Z1
1040 H6(I)=F(Q1(I))
1050 NEXT I
```

```
1060 PRINT @37,26:1
1070 PRINT @4Ø:" C      T      F"
1080 DIM U(3)
1090 FOR I=1 TO Z1
1100 U(1)=F6(I)
1110 U(2)=G6(I)
1120 U(3)=H6(I)
1130 PRINT @4Ø: USING 115Ø:U
1140 NEXT I
1150 IMAGE 3(2d,4x)
1160 PRINT @37,26:Ø
1170 END
1180 PRINT "WHICH FILE?"
1190 INPUT F9
1200 FIND F9
1210 DELETE QØ
1220 DIM QØ(4,125)
1230 ON EOF (Ø) THEN 127Ø
1240 FOR I=1 TO 125
1250 INPUT @33:QØ(4,I),S1,QØ(3,I),QØ(2,I),QØ(1,I)
1260 NEXT I
1270 I9=I-1
1280 DELETE HØ,H1
1290 DIM HØ(2,1ØØ),H1(2,1ØØ)
1300 J1=1
1310 J2=1
1320 K1=1
1330 K2=1
1340 FOR I=1 TO I9
1350 IF QØ(1,I)=Ø THEN 142Ø
1360 NEXT I
1370 K1=K1-1
1380 K2=K2-1
1390 J1=J1-1
1400 J2=J2-1
```

```
1410 GO TO 162ø
1420 IF Qø(2,I)<4 THEN 144ø
1430 GO TO 152ø
1440 IF Qø(3,I)=1ø THEN 136ø
1450 IF Qø(3,I)=11 THEN 149ø
1460 Hø(2,J1)=Qø(4,I)
1470 J1=J1+1
1480 GO TO 136ø
1490 Hø(1,J2)=Qø(4,I)
1500 J2=J2+1
1510 GO TO 136ø
1520 IF Qø(3,I)=1ø THEN 136ø
1530 IF Qø(3,I)=11 THEN 157ø
1540 H1(2,K1)=Qø(4,I)
1550 K1=K1+1
1560 GO TO 136ø
1570 H1(1,K2)=Qø(4,I)
1580 K2=K2+1
1590 GO TO 136ø
1600 REM:RAMDOM SAMPLE--ORDER--11-HOMO;12-Homo
1610 REM:11-HETER;12-HETER
1620 PRINT "SAMPLE SIZE?"
1630 INPUT Z1
1640 DIM R5(Z1),S5(Z1),U5(Z1),W5(Z1)
1650 REM Hø(1) = NO EXP:HOMO:11
1660 REM Hø(2)=NO EXP:HOMO:12
1670 REM H1(1)= NO EXP:HETER:11
1680 REM H1(2)=NO EXP: HETER:12
1690 Z2=J2
1700 GOSUB 253ø
1710 FOR I=1 TO Z1
1720 R5(I)=Hø(1,Q1(I))
1730 NEXT I
1740 Z2=J1
1750 GOSUB 253ø
```

```
1760 FOR I=1 TO Z1
1770 S5(I)=HØ(2,Q1(I))
1780 NEXT I
1790 Z2=K2
1800 GOSUB 253Ø
1810 FOR I=1 TO Z1
1820 U5(I)=H1(1,Q1(I))
1830 NEXT I
1840 Z2=K1
1850 GOSUB 253Ø
1860 FOR I=1 TO Z1
1870 W5(I)=H1(2,Q1(I))
1880 NEXT I
1890 PRINT @37,26:1
1900 PRINT @4Ø:"11-HOMO 12-HOMO 11-HET 12-HET"
1910 DIM W7(4)
1920 FOR I=1 TO Z1
1930 W7(1)=R5(I)
1940 W7(2)=S5(I)
1950 W7(3)=U5(I)
1960 W7(4)=W5(I)
1970 PRINT @4Ø: USING 199Ø:W7
1980 NEXT I
1990 IMAGE 4(2D,6X)
2000 PRINT @37,26:Ø
2010 END
2020 PRINT "WHICH FILE?"
2030 INPUT F9
2040 FIND F9
2050 DELETE QØ
2060 DIM QØ(2,8Ø)
2070 ON EOF (Ø) THEN 211Ø
2080 FOR I=1 TO 8Ø
2090 INPUT @33:QØ(2,I),QØ(1,I),GØ,G1,SØ
2100 NEXT I
```

```
2110 I9=I-1
2120 DELETE LØ
2130 DIM LØ(2,4Ø)
2140 J1=1
2150 J2=1
2160 FOR I=1 TO I9
2170 IF QØ(1,I)=2 THEN 226Ø
2180 IF QØ(1,I)=6 THEN 223Ø
2190 NEXT I
2200 J1=J1-1
2210 J2=J2-1
2220 GO TO 231Ø
2230 LØ(1,J2)=QØ(2,I)
2240 J2=J2+1
2250 GO TO 219Ø
2260 LØ(2,J1)=QØ(2,I)
2270 J1=J1+1
2280 GO TO 219Ø
2290 STOP
2300 REM RANDOM SAMPLE --ORDER-WMHS --MCJHS
2310 PRINT "SAMPLE SIZE?"
2320 INPUT Z1
2330 DELETE P5,R5
2340 DIM P5(Z1),R5(Z1)
2350 Z2=J2
2360 GOSUB 253Ø
2370 FOR I=1 TO Z1
2380 P5(I)=LØ(1,Q1(I))
2390 NEXT I
2400 Z2=J1
2410 GOSUB 253Ø
2420 FOR I=1 TO Z1
2430 R5(I)=LØ(2,Q1(I))
2440 NEXT I
2450 PRINT @37,26:1
```

```
2460 PRINT @4Ø:"GRP EXP     NO GRP EXP"
2470 FOR I=1 TO Z1
2480 PRINT @4Ø: USING 25ØØ:R5(I),P5(I)
2490 NEXT I
2500 IMAGE 2(2D,6X)
2510 PRINT @37,26:Ø
2520 END
2530 REM RANDOM SELECTION W/O REPLACEMENT
2540 DELETE Q1
2550 DIM Q1(Z1)
2560 FOR I=1 TO Z1
2570 Q1(I)=INT ( Z2*RND(-2))+1
2580 NEXT I
2590 FOR I=1 TO Z1-1
2600 FOR J=I+1 TO Z1
2610 IF Q1(I)=Q1(J) THEN 265Ø
2620 NEXT J
2630 NEXT I
2640 RETURN
2650 Q1(J)=INT(Z2*RND(-2))+1
2660 GO TO 259Ø
2670 PRINT "WHICH FILE?"
2680 INPUT F9
2690 FIND F9
2700 DELETE QØ
2710 DIM QØ(1ØØ)
2720 ON EOF (Ø) THEN 276Ø
2730 FOR I=1 TO 1ØØ
2740 INPUT @33:QØ(I),S1,GØ,G1,SØ
2750 NEXT I
2760 I9=I-1
2770 Z2=I9
2780 PRINT "Sample Size?"
2790 INPUT Z1
2800 GOSUB 253Ø
```

```
2810 DELETE W4
2820 DIM W4(Z1)
2830 FOR I=1 TO Z1
2840 W4(I)=QØ(Q1(I))
2850 NEXT I
2860 PRINT @37,26:1
2870 PRINT @4Ø:"GROUP"
2880 FOR I=1 TO Z1
2890 PRINT @4Ø: USING 291Ø:W4(I)
2900 NEXT I
2910 IMAGE 3D
2920 PRINT @37,26:Ø
2930 END
2940 REM  selects heter-homo no grp exp,
2950 PRINT "which file?"
2960 INPUT F9
2970 FIND F9
2980 DELETE QØ
2990 DIM QØ(3,1ØØ)
3000 ON EOF (Ø) THEN 3Ø4Ø
3010 FOR I=1 TO 1ØØ
3020 INPUT @33:QØ(1,I),S1,GØ,QØ(2,I),QØ(3,I)
3030 NEXT I
3040 I9=I-1
3050 DIM H1(1ØØ),HØ(1ØØ)
3060 J1=1
3070 J2=1
3080 FOR I=1 TO I9
3090 IF QØ(3,I)=Ø THEN 314Ø
3100 NEXT I
3110 J1=J1-1
3120 J2=J2-1
3130 GO TO 322Ø
3140 IF QØ(2,I)<4 THEN 318Ø
3150 HØ(J2)=QØ(1,I)
```

```
3160 J2=J2+1
3170 GO TO 31ØØ
3180 H1(J1)=QØ(1,I)
3190 J1=J1+1
3200 GO TO 31ØØ
3210 REM  random samples order hom-hetero
3220 PRINT "Sample Size?"
3230 INPUT Z1
3240 DIM L6(Z1),L7(Z1)
3250 REM  hØ(1  )=no exp-hetero
3260 REM  h1(1  )=no exp-homo
3270 Z2=J1
3280 GOSUB 254Ø
3290 FOR I=1 TO Z1
3300 L6(I)=H1(Q1(I))
3310 NEXT I
3320 Z2=J2
3330 GOSUB 254Ø
3340 FOR I=1 TO Z1
3350 L7(I)=HØ(Q1(I))
3360 NEXT I
3370 PRINT @37,26:1
3380 PRINT @4Ø:"homo  hetero"
3390 DIM B3(2)
3400 FOR I=1 TO Z1
3410 B3(1)=L6(I)
3420 B3(2)=L7(I)
3430 PRINT @4Ø: USING 345Ø:B3
3440 NEXT I
3450 IMAGE 2(2d,6x)
3460 PRINT @37,26:Ø
3470 END
```

Calculation of CAP Measure of Intellectual Development and Frequency Distribution

```
1 PRINT "ro's  research "
2 PRINT "press 1  then 2  then 3  BREAK Twice at end then 4
  good luck"
3 GO TO 19Ø
4 PRINT @37,26:1
5 GO TO 68Ø
8 PRINT "which tape file?"
9 INPUT F
10 FIND F
11 RETURN
12 GO TO 35Ø
13 GO TO 68Ø
14 FIND F
15 RETURN
16 CLOSE
17 RETURN
20 DELETE A$
21 DIM A$(72)
22 GO TO 79Ø
24 GOSUB 8
25 DELETE A$
26 GO TO 118Ø
100 REM Variables are defined as follows:
110 REM eØ=EFT, S=S, F=F, G=G, SØ=sex(1=male,2=female)
120 REM gØ=grade, aØ=age in months, sl=school (2=WMH, 3=MHS, 4=SHS,
130 REM  5=CIHS, 6=CJHS), A$=student's name, al=student's number,
140 REM  e=EI score, gl=group type (1=homogeneous, 2=heterogeneous)
150 REM  g2=group number
160 REM  standard tape format is al,a$,sl,gØ,aØ,sØ,eØ,s,f,g,
    e,gl,g2
170 REM in image 4d,2x,24A,2x,ld,2x,2d,2x,3d,2x,ld,2x,2d,2x,ld,2x,
180 REM  ld,2x,ld,2x,2d.2d,2x,3d,2x,3d
190 B$="STUDENT'S NUMBER?"
200 DIM A$(24)
210 C$="NAME?"
```

```
220 D$="EFT?"
230 E$="S?"
240 F$="F?"
250 G$="G?"
260 H$="AGE(MONTHS)?"
270 I$="SEX?"
280 J$="SCHOOL?"
290 K$="GROUP TYPE?"
300 L$="GROUP NUMBER?"
310 M$="GRADE?"
320 RETURN
330 IMAGE 4D,2X,24A,2X,1D,2X,2D,2X,3D,2X,1D,2X,2D,2X,1D,2X,
    1D,2X,1D,2X
340 IMAGE 2d.2d,2x,3d,2x,3d
350 PAGE
360 DIM N$(1)
370 N$=" "
380 PRINT B$
390 INPUT A1
400 PRINT C$
410 INPUT A$
420 PRINT D$
430 INPUT EØ
440 PRINT E$
450 INPUT S
460 PRINT F$
470 INPUT F
480 PRINT G$
490 INPUT G
500 PRINT H$
510 INPUT AØ
520 PRINT I$
530 INPUT SØ
540 PRINT J$
550 INPUT S1
```

```
560 PRINT K$
570 INPUT G1
580 PRINT L$
590 INPUT G2
600 PRINT M$
610 INPUT GØ
620 E=Ø.17*EØ+Ø.38*S+Ø.37*F+Ø.3*G+3.95
630 GOSUB 71Ø
640 GOSUB 76Ø
650 PRINT " "
660 GO TO 35Ø
670 RETURN
680 PRI @4Ø:"St. #        Name          EFT    S   F   G    EI    IL   GD"
690 PRINT @37,26:Ø
700 RETURN
710 PRINT @37,26:1
720 PRINT @4Ø: USING 73Ø:A1,A$,EØ,S,F,G,E,N$,GØ
730 IMAGE 4d,4x,24a,2d,3x,1d,2x,1d,2x,1d,2x,2d.2d,2X,1A,2X,2D
740 PRINT @37,26:Ø
750 RETURN
760 PRINT @33: USING 33Ø:A1,A$,S1,GØ,AØ,SØ,EØ,S,F,G;
770 PRINT @33: USING 34Ø:E;G1;G2
780 RETURN
790 GOSUB 8
800 PRINT @37,26:1
810 GOSUB 68Ø
820 GOSUB 95Ø
830 REM N$=INTELLECTUAL LEVEL
840 E=Ø.17*EØ+Ø.38*S+Ø.37*F+Ø.3*G+3.95
850 REM E MAY BE UNRELIABLE RECALCULATE  E -- S,FAND G ARE RELIABLE
860 IF E=>11.5 THEN 9ØØ
870 IF E=>8.5 THEN 92Ø
880 N$="C"
890 GO TO 93Ø
900 N$="F"
```

```
910 GO TO 93Ø
920 N$="T"
930 GOSUB 71Ø
940 GO TO 82Ø
950 INPUT @33:A$
960 INPUT @33:E,G1,G2
970 B$=SEG(A$,1,4)
980 C$=SEG(A$,7,24)
990 D$=SEG(A$,32,3)
1000 E$=SEG(A$,36,4)
1010 F$=SEG(A$,4Ø,3)
1020 G$=SEG(A$,44,3)
1030 H$=SEG(A$,48,3)
1040 I$=SEG(A$,52,2)
1050 J$=SEG(A$,55,2)
1060 K$=SEG(A$,58,3)
1070 A1=VAL(B$)
1080 S1=VAL(D$)
1090 GØ=VAL(E$)
1100 AØ=VAL(F$)
1110 SØ=VAL(G$)
1120 EØ=VAL(H$)
1130 S=VAL(I$)
1140 F=VAL(J$)
1150 G=VAL(K$)
1160 A$=C$
1170 RETURN
1180 ON EOF (Ø) THEN 135Ø
1190 PRINT "grade?"
1200 INPUT YØ
1210 A9=Ø
1220 A8=Ø
1230 A7=Ø
1240 GOSUB 95Ø
1250 IF GØ=YØ THEN 127Ø
```

```
1260 GO TO 124Ø
1270 IF E=>11.5 THEN 131Ø
1280 IF E=>8.5 THEN 133Ø
1290 A7=A7+1
1300 GO TO 124Ø
1310 A9=A9+1
1320 GO TO 124Ø
1330 A8=A8+1
1340 GO TO 124Ø
1350 PRINT @37,26:1
1360 PRINT @4Ø:"grade     f    t     c"
1370 PRINT @4Ø: USING "4(2x,3d)":YØ,A9,A8,A7
1380 PRINT @37,26:Ø
1390 END
```

<u>Calculation of Correlation between Scores on the Piagetian Task Interviews and the CAP Measure of Intellectual Development</u>

```
1 PRINT "ro's  research "
2 PRINT "press 1  then 2  then 3  BREAK Twice at end then 4
  good luck"
3 GO TO 19Ø
4 PRINT @37,26:1
5 GO TO 56Ø
8 PRINT "which tape file?"
9 INPUT F
10 FIND F
11 RETURN
12 GO TO 31Ø
16 CLOSE
17 RETURN
20 GO TO 67Ø
21 REM READ DATA OFF TAPE
24 GO TO 97Ø
25 REM SORTING STDS INTO WHOLE GRP;C;T;F
28 GO TO 79Ø
```

```
29 REM RANDOM # GENERATION
36 GO TO 118Ø
37 REM SELECT RANDOM SAMPLES & CALCULATE R:WHOLE GRP
40 GO TO 126Ø
41 REM SELECT RANDOM SAMPLES & CALCULATE R:C
44 GO TO 134Ø
45 REM SELECT RANDOM SAMPLES & CALCULATE R:T
48 GO TO 142Ø
49 REM SELECT RANDOM SAMPLES & CALCULATE R:F
100 REM Variables are defined as follows:
110 REM eØ=EFT, S=S, F=F, G=G, SØ=CONS OF VOL
120 REM gØ=PIAGET SCORE,sl=sEP oF VAR
130 REM al=student's number,
140 REM  e=EI score, gl=EQUIL IN BAL
150 REM  g2=COL LIQ
160 REM  standard tape format is al,sl,gØ,sØ,eØ,s,f,g,e,gl,g2
170 REM  in image 4d,2x,ld,2x,2d,2x,ld,2x,2d,2x,ld,2x,
180 REM  ld,2x,ld,2x,2d.2d,2x,ld,2x,ld
190 B$="STUDENT'S NUMBER?"
200 D$="EFT?"
210 E$="S?"
220 F$="F?"
230 G$="G?"
240 H$="CONS OF VOL?"
250 I$="SEP OF VAR?"
260 J$="EQUIL IN BAL?"
270 K$="COL LIQ"
280 RETURN
290 IMAGE 4D,2X,1D,2X,2D,2X,1D,2X,2D,2X,1D,2X,1D,2X,1D
300 IMAGE 2d.2d,2x,ld,2x,ld
310 PAGE
320 PRINT B$
330 INPUT A1
340 PRINT D$
350 INPUT EØ
```

```
360 PRINT E$
370 INPUT S
380 PRINT F$
390 INPUT F
400 PRINT G$
410 INPUT G
420 PRINT H$
430 INPUT SØ
440 PRINT I$
450 INPUT S1
460 PRINT J$
470 INPUT G1
480 PRINT K$
490 INPUT G2
500 E=Ø.17*EØ+Ø.38*S+Ø.37*F+Ø.3*G+3.95
510 GØ=SØ+S1+G1+G2
520 GOSUB 59Ø
530 GOSUB 64Ø
540 PRINT " "
550 GO TO 31Ø
560 PRI @4Ø:"ST.# C.VOL S.VAR E.BAL C.LIQ PIA.SC  EFT
    S    F    G    EI"
570 PRINT @37,26:Ø
580 RETURN
590 PRINT @37,26:1
600 PRINT @4Ø: USING 61Ø:A1,SØ,S1,G1,G2,GØ,EØ,S,F,G,E
610 IMA4d,3x,1D,5X,1D,4X,1D,5X,1D,6X,2D,6X,2D,4x,1d,3x,1d,
    3x,1d,3x,2d.2d
620 PRINT @37,26:Ø
630 RETURN
640 PRINT @33: USING 29Ø:A1,S1,GØ,SØ,EØ,S,F,G
650 PRINT @33: USING 3ØØ:E;G1;G2
660 RETURN
670 REM SELECTION SUBROUTINE
680 PRINT "WHICH FILE?"
690 INPUT F
```

```
700 FIND F
710 ON EOF (Ø) THEN 76Ø
720 DIM QØ(2,4ØØ)
730 FOR I=1 TO 4ØØ
740 INPUT @33:A1,S1,QØ(1,I),SØ,EØ,S,F,G,QØ(2,I),G1,G2
750 NEXT I
760 I9=I-1
770 END
780 REM RANDOM # GENERATION
790 PRINT "INPUT SAMPLE SIZE"
800 INPUT Z
810 PRINT "INPUT RANGE OF POPULATION"
820 INPUT Z8
830 DELETE Q1
840 DIM Q1(Z)
850 FOR I=1 TO Z
860 Q1(I)=INT(Z8*RND(-2))+1
870 NEXT I
880 FOR I=1 TO Z-1
890 FOR J=I+1 TO Z
900 IF Q1(I)=Q1(J) THEN 94Ø
910 NEXT J
920 NEXT I
930 RETURN
940 Q1(J)=INT(Z8*RND(-2))+1
950 GO TO 88Ø
960 REM PIAGETIAN SCORE SELECTION PRIOR TO CORRELATION FOR
    I=1 TO Z
970 DIM S7(2,2ØØ),S8(2,2ØØ),S9(2,2ØØ)
980 P7=1
990 P8=1
1000 P9=1
1010 FOR I=1 TO I9
1020 IF QØ(1,I)<=8 THEN 1Ø9Ø
1030 IF QØ(1,I)<=11 THEN 113Ø
```

```
1040 S9(1,P9)=QØ(1,I)
1050 S9(2,P9)=QØ(2,I)
1060 P9=P9+1
1070 NEXT I
1080 RETURN
1090 S7(1,P7)=QØ(1,I)
1100 S7(2,P7)=QØ(2,I)
1110 P7=P7+1
1120 GO TO 1Ø7Ø
1130 S8(1,P8)=QØ(1,I)
1140 S8(2,P8)=QØ(2,I)
1150 P8=P8+1
1160 GO TO 1Ø7Ø
1170 REM SELECTION RANDOM STUDENTS --WHOLE GROUP;C;T;F
1180 DELETE X,Y,Z1,Z2
1190 DIM X(2ØØ),Y(2ØØ),Z1(2ØØ),Z2(2ØØ)
1200 FOR I=1 TO Z
1210 X(I)=QØ(1,Q1(I))
1220 Y(I)=QØ(2,Q1(I))
1230 NEXT I
1240 GOSUB 151Ø
1250 RETURN
1260 DELETE X,Y,Z1,Z2
1270 DIM X(2ØØ),Y(2ØØ),Z1(2ØØ),Z2(2ØØ)
1280 FOR I=1 TO Z
1290 X(I)=S7(1,Q1(I))
1300 Y(I)=S7(2,Q1(I))
1310 NEXT I
1320 GOSUB 151Ø
1330 RETURN
1340 DELETE X,Y,Z1,Z2
1350 DIM X(2ØØ),Y(2ØØ),Z1(2ØØ),Z2(2ØØ)
1360 FOR I=1 TO Z
1370 X(I)=S8(1,Q1(I))
1380 Y(I)=S8(2,Q1(I))
```

```
1390 NEXT I
1400 GOSUB 1510
1410 RETURN
1420 DELETE X,Y,Z1,Z2
1430 DIM X(200),Y(200),Z1(200),Z2(200)
1440 FOR I=1 TO Z
1450 X(I)=S9(1,Q1(I))
1460 Y(I)=S9(2,Q1(I))
1470 NEXT I
1480 GOSUB 1510
1490 RETURN
1500 REM CORRELATION SUBROUTINE
1510 X1=0
1520 Y1=0
1530 M1=0
1540 M2=0
1550 Q=Z
1560 FOR I=1 TO Q
1570 X1=X1+X(I)*X(I)
1580 Y1=Y1+Y(I)*Y(I)
1590 M1=X(I)+M1
1600 M2=Y(I)+M2
1610 NEXT I
1620 M1=M1/Q
1630 M2=M2/Q
1640 S1=SQR(X1/Q-M1*M1)
1650 S2=SQR(Y1/Q-M2*M2)
1660 FOR I=1 TO Q
1670 Z1(I)=(X(I)-M1)/S1
1680 Z2(I)=(Y(I)-M2)/S2
1690 NEXT I
1700 R=0
1710 FOR I=1 TO Q
1720 R=R+Z1(I)*Z2(I)
1730 NEXT I
```

```
1740 R=R/Q
1750 T8=SQR(1-R*R)
1760 PRINT @37,26:1
1770 PRINT @40:"DATA & MEANS"
1780 FOR I=1 TO Q
1790 PRINT @40: USING 1810:X(I),Y(I)
1800 NEXT I
1810 IMAGE 3D,2X,3D
1820 PRINT @37,26:1
1830 PRINT @40:"X MEAN         Y MEAN"
1840 PRINT @40: USING 1850:M1,M2
1850 IMAGE 3D.3D,10X,3D.3D
1860 PRINT @37,26:1
1870 PRINT @40:""
1880 PRINT @40:"X STD DEV          Y STD DEV"
1890 PRINT @40: USING 1900:S1,S2
1900 IMAGE 3D.3D,15X,3D.3D
1910 PRINT @37,26:1
1920 PRINT @40:""
1930 PRINT @40:"R VALUE"
1940 PRINT @40:R
1950 PRINT @37,26:1
1960 PRINT @40:""
1970 PRINT @40:"STD ERROR OF ESTIMATE"
1980 PRINT @40:T8
1990 PRINT @37,26:0
2000 RETURN
```

t-Test (Equal or Unequal Cell Sizes)

```
1 REM T TEST
4 PAGE
5 GO TO 100
100 PRINT "FIRST GROUP SIZE?"
110 INPUT N
120 PRINT "SECOND GROUP SIZE?"
```

```
130 INPUT M
135 DELETE X,G,Y,H
140 DIM X(N),G(N),Y(M),H(M)
150 FOR I=1 TO N
160 INPUT X(I)
170 PRINT @37,26:1
180 PRINT @4Ø:"  ";X(I)
190 PRINT @37,26:Ø
200 NEXT I
210 PRINT "GGGGGGG"
215 PRINT @4Ø: USING 216:
216 IMAGE 8L
220 FOR I=1 TO M
230 INPUT Y(I)
240 PRINT @37,26:1
250 PRINT @4Ø:"  ";Y(I)
260 PRINT @37,26:Ø
270 NEXT I
280 S=Ø
290 GØ=Ø
300 FOR I=1 TO N
310 S=S+X(I)
320 G(I)=X(I)*X(I)
330 GØ=GØ+G(I)
340 NEXT I
350 U=(N*GØ-S*S)/(N*(N-1))
360 XØ=S/N
370 S=Ø
380 HØ=Ø
390 FOR I=1 TO M
400 S=S+Y(I)
410 H(I)=Y(I)*Y(I)
420 HØ=HØ+H(I)
430 NEXT I
440 V=(M*HØ-S*S)/(M*(M-1))
```

```
450 YØ=S/M
460 E=1/N+1/M
470 Q=N+M-2
480 T=(XØ-YØ)/SQR(((N-1)*U+(M-1)*V)/Q*E)
490 A$=" MEAN 1"
500 B$="MEAN 2"
510 C$="VARIANCE 1"
520 D$="VARIANCE 2"
530 E$="   T "
540 F$="D.F."
550 PRINT  @37,26:1
560 PRINT @4Ø: USING 57Ø:A$,B$,C$,D$,E$,F$
570 IMAGE 6(12A)
580 PRINT @4Ø: USING 59Ø:XØ,YØ,U,V,T,Q
590 IMAGE 3D.3D,4X,3D.3D,4X,5D.3D,4X,5D.3D,4X,3D.3D,4X,4D
600 PRINT @37,26:Ø
610 PRINT @4Ø: USING 216:
```

<u>Generating Random Samples without Replacement</u>

```
4 GO TO 1ØØ
100 PRINT "INPUT POP RANGE"
110 INPUT Z8
120 PRINT "INPUT SAMPLE SIZE"
130 INPUT Z
135 DELETE Q1
150 DIM Q1(Z)
160 FOR I=1 TO Z
170 Q1(I)=INT(Z8*RND(-2))+1
180 NEXT I
190 FOR I=1 TO Z-1
200 FOR J=I+1 TO Z
210 IF Q1(I)=Q1(J) THEN 24Ø
220 NEXT J
230 NEXT I
235 GO TO 26Ø
```

```
240 Q1(J)=INT(Z8*RND(-2))+1
250 GO TO 19Ø
260 FOR I=1 TO Z
265 PRINT @37,26:1
270 PRINT @4Ø:Q1(I)
280 NEXT I
285 PRINT @4Ø: USING 31Ø:
290 PRINT @37,26:Ø
300 END
310 IMAGE 6L
```

APPENDIX I

APPENDIX I

NON-ORTHOGONAL ANOVA'S

A non-orthogonal ANOVA refers to a design in which the number of observations are not equal in every cell. The usual ANOVA methods of analysis are not always recommended for analyzing non-orthogonal designs because the assumption of equality of population variances may be violated in such a way that the tests are invalid. Several recommendations for treating non-orthogonal designs will be discussed.

Winer (1971, pp. 402-422) discusses two methods for non-orthogonal designs. The first method uses an unweighted means approach. This method is used when the unequal cell sizes arise because of some chance factor in the experiment. Effectively this method considers that these factors which caused the unequal cell sizes have nothing to do with the significance of the experiment. The disadvantage of the unweighted means approach is that it does not minimize the sum of least squares. The least squares is the second approach which is recommended to be used when something systematic caused the cells to be unequal. This approach places more emphasis on cells with larger sizes and regards these cells as more important than smaller cell sizes. The disadvantage of the least squares approach is that is is computationally difficult.

Model comparison methods as suggested by Applebaum and Cramer (1974, pp. 335-343) have problems different to those methods discussed by Winer (1971). The major problem is that there are several different ways of computing the required F-ratio, but each of these ways has a different meaning. Applebaum and Cramer (1974) not only recommend the model comparison procedure but also suggest that the interaction test should be carried out first, before the tests with the independent variables. If the interaction test is significant, Applebaum and Cramer (1974) suggest that no further tests are necessary. Only if the interaction test is non-significant should the other tests be carried out. There is some controversy over this recommendation because the interaction is a unique effect due to the combined effect of the independent variables. There are some circumstances when it is possible to interpret the main effect (or independent variable) results even though the interaction effect is significant. One further problem arises with Applebaum and Cramer's recommendation. If the interaction effect is ignored in both models, the result is that the denominators of the F-ratios include sums of squares and the degrees of freedom of the interaction effect as well as the within variability effect. Consequently, under these circumstances a conservative F-test would result. It could be argued that before pooling is carried out the interaction effects should be zero, not just small.

Another approach is discussed by Herr and Gaebelein (1978, pp. 207-216). They point out that each analysis for the row and column tests have their own particular meaning and they suggest that the researcher should select the approach which is most appropriate in achieving the goal of the experiment. Explicit statements are given for the hypotheses together with their appropriate model comparisons. Three different approaches were given for each row and column hypothesis. For the first approach the main effects were defined in terms of the unweighted averages of the cell means. Weighted averages of the cell means are used for the second approach. Parametric statements of hypotheses using the weighted averages are difficult to interpret. The hypothesis associated with the row effect may, however, be interpreted as whether or not this effect explains a significant amount of the variability if the column and interaction effects are ignored. A similar hypothesis is proposed for the column effect. Rather than using the ignoring tests of the second approach, the row and column hypotheses in the third approach are tested by eliminating effects. For example, the row effect hypothesis is concerned with whether or not this effect explains a significant amount of the variability in the presence of, or adjusting for, column effects and ignoring interaction effects. A similar interpretation is given for the hypothesis associated with the column effect. The third approach also uses weighted averages.

Although the averages are different to those in the second approach the parametric statements of the third approach hypotheses are just as difficult to interpret. All three approaches are equivalent for orthogonal designs and lead to the same results.

If a non-orthogonal design is to be analyzed, Herr and Gaebelein (1978) suggest that the researcher determine which approach and interpretation are most relevant to the experiment. Consideration should also be given to the properties of the analysis which follow the interpretation. Once a particular hypothesis is chosen there can be only one analysis. The most important property to be considered is the sums of squares and whether or not these are uncorrelated or independent. Uncorrelated sums of squares are also orthogonal. If the sums of squares are non-orthogonal, then the tests have reduced power and sources of variability may be hidden or lost by the analytical procedures. These effects of non-orthogonality depend on the cell sizes and the arrangement of the unequal cell sizes.

Herr and Gaebelein (1978) combine the three approaches for interpreting hypotheses discussed above into five different types of tests. Of these five tests only two gave completely orthogonal analyses in which the row, column and interaction sums of squares were orthogonal with each other and with the total sums of squares. Therefore, these two tests would lead a test with the maximum power and a decrease in the probability of a Type I error. The two tests were designated as hierarchical with columns first and rows adjusted for columns (HCR) and hierarchical with rows first and columns adjusted for rows (HRC). In the HCR test the column effect was tested with the second approach and the row effect test with the third approach. The HRC test involved the same approaches as the HCR test except those approaches were used in the reverse order--rows tested with the second approach, columns tested with the third approach.

Different methods of treating non-orthogonal designs have also been discussed by Speed, Hocking and Hackney (1978, pp. 105-111). These authors also suggest that the selection of a method of analysis should be based on the appropriateness of the hypotheses being investigated rather than upon heuristic or ease of computation grounds. Furthermore, the authors suggest that orthogonality or independence is insufficient justification for a method if the hypotheses from which that method arises have no meaningful interpretation. Hypotheses which depend on particular cell frequencies are also difficult to justify.

Of the eight general methods of analysis for non-orthogonal designs discussed by Speed, Hocking and Hackney (1978), the method by Overall and Spiegel which deals with a priori ordering corresponds to the HRC and HCR tests discussed by Herr and Gaebelein (1978). Although the parametric statements for the hypotheses used in the HCR and HRC tests are difficult to interpret, these tests were used in this research. The interpretations of these hypotheses given by Herr and Gaebelein (1978) were accepted for this research. It was then possible to have tests which gave maximum power and in which sources of variability were not lost or hidden by the analytical procedures. Contrary to the recommendations of Speed, Hocking and Hackney (1978) orthogonality was a major concern in this research. This was justified on the grounds that the interpretations (whether or not a main effect explains a significant amount of variability, ignoring interaction and either ignoring or eliminating the other main effects) for the HCR and HRC tests given by Herr and Gaebelein were meaningful for this research. As recommended by these authors, the Statistical Package for the Social Sciences (Nie, et al., 1975) Option ten was used to test the non-orthogonal ANOVA's in this research.

References

Applebaum, M. I. and Cramer, E. M. Some Problems in the Non-Orthogonal Analysis of Variance. Psychological Bulletin, 1974, 81, No. 6.

Herr, D. G. and Gaebelein, J. Nonorthogonal Two-Way Analysis of Variance. Psychological Bulletin, 1978, 85, No. 1.

Nie, N. H. et al. Statistical Package for the Social Sciences. New York: McGraw-Hill Book Company, 1975.

Speed, F. M., Hocking, R. R. and Hackney, O. P. Methods of Analysis of Linear Models with Unbalanced Data. Journal of the American Statistical Association, March 1978, 73, No. 361.

Winer, B. J. Statistical Principles in Experimental Design, 2nd ed. New York: McGraw-Hill Book Company, 1971.

APPENDIX J

APPENDIX J

EQUATIONS FOR THE NON-CENTRALITY PARAMETER, ϕ USED IN CALCULATING THE POWER OF TESTS

Two-Way ANOVA

Equal Cell Sizes

$$\phi_A = \sqrt{\frac{IK\Sigma\alpha_j^2}{J\sigma_e^2}}$$

Unequal Cell Sizes

$$\phi_A = \sqrt{\frac{\sum_j (\sum_K n_{jk})\alpha_j^2}{J\sigma_e^2}}$$

One-Way ANOVA

Equal Cell Sizes

$$\phi_A = \sqrt{\frac{I\Sigma\alpha_j^2}{J\sigma_e^2}}$$

Unequal Cell Sizes

$$\phi_A = \sqrt{\frac{\sum_j n_j\alpha_j^2}{J\sigma_e^2}}$$

APPENDIX K

APPENDIX K

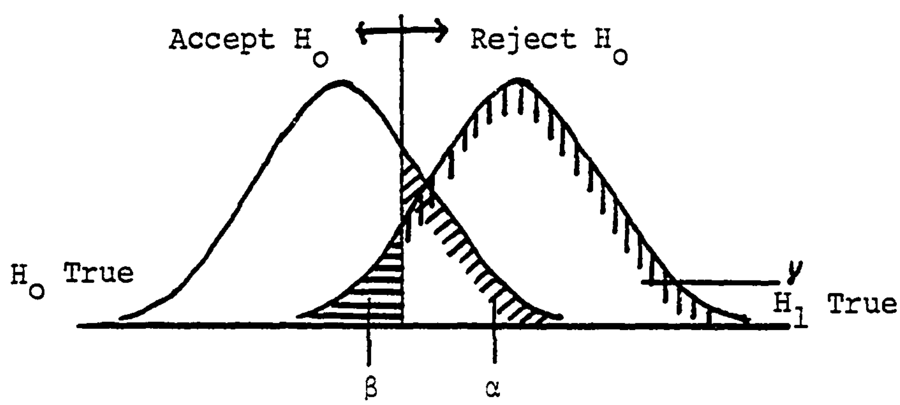
STATISTICAL CONSIDERATIONS (TYPE I AND TYPE II ERRORS, POWER OF TESTS)

A Type I error can be controlled by setting the level of significance for a research study at a very low level. The level of significance, α, is the probability of erroneously rejecting the null hypothesis. Type II errors, however, cannot be directly controlled by the investigator. Nevertheless, the probability of a Type II error is related to the level of significance and the power of the statistical test.

Sampling distributions for the true, null and alternative hypotheses are shown in Graph III.

GRAPH III

SAMPLING DISTRIBUTIONS FOR H_0 TRUE AND H_1 TRUE



It can be seen that if the probability, α, of falsely rejecting H_o increases, the area, β, under the true alternative decreases. The probability of a Type II error is represented by β. If β decreases, the area marked γ under the true alternative, H_1 increases. This area represents the probability of correctly rejecting the null hypothesis. The ability of a test to detect differences is the power of the test. In other words, the power of a test is the probability of correctly rejecting H_o given a particular value of the population parameter. Hence, if the power of the test is sufficiently large, the probability of a Type II error is small.

Power can be regulated because it is related to sample size, level of significance, α, and population error variance, σ_e^2, through the non-centrality parameter, ϕ. This relationship is shown in equation 1. See Appendix J for equations of ϕ using unequal cell sizes.

$$\phi_A = \sqrt{\frac{IKL\ \Sigma\alpha_j^2}{J\sigma_e^2}} \qquad (1)$$

ϕ_A = non-centrality parameter for the A main effect

I = common cell size

J = # A main effect levels

K = # B main effect levels

L = # C main effect levels

σ_e^2 = population error variance

α_j = effect of treatment

Graphs showing the relationship between ϕ_A, α and power are given by Pearson and Hartley (1951). Furthermore, the power is affected by the type of test. For the same α and true alternative, one-tailed tests are more powerful if the true alternative is in the direction of the rejection region.

A more complete description of the factors which affect the power of a test is given in Hays (1973, pp. 357-369).

For the purposes of this research, cell sizes were calculated which achieved a power of 95 per cent at a level of significance of $\alpha = 0.05$ and population error standard deviation differences of 1.0 σ_e and 1.5 σ_e. The Pearson and Hartley (1951) charts have only been prepared for the 0.01 and 0.05 levels of significance. Most of the tests in this research required minimum probabilities of Type II error and these probabilities decrease with increasing levels of significance. Therefore, the cell sizes were calculated using the highest level of significance available on the charts--$\alpha = 0.05$. As explained in Chapter IV, sufficient data for these calculated cell sizes were not always possible and the tests were carried out with the available data. Under these circumstances, the powers of the tests were calculated for selected standard deviation differences and the cell sizes used in the tests.

References

Hays, W. L. Statistics for the Social Sciences, 2nd ed. New York: Holt, Rinehart and Winston, Inc., 1973.

Pearson, E. S. and Hartley, H. O. Charts of the Power Function for Analysis of Variance Tests Derived from the Non-Central F-Distribution. Biometrika, 1951, 38, pp. 112-130.